

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 23:49:13 ; Search time 534.111 Seconds
(without alignments)
6894.813 Million cell updates/sec

Title: US-09-551-494-2_COPY_5443_5518

Perfect score: 76

Sequence: 1 gtgagagacgaggcccat.....agatgtccctatgtcgatca 76

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pet.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ets.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	76	100.0	396	14	TMO307579	AJ307579 Tobacco m
2	76	100.0	632	6	A68760	A68760 Sequence 1
3	76	100.0	706	14	TMO308689	AJ308689 Tobacco m
4	76	100.0	750	14	TMO309080	AJ509080 Tobacco m
5	76	100.0	750	14	TMO309084	AJ509084 Tobacco m
6	76	100.0	782	14	TMO308692	AJ308692 Tobacco m
7	76	100.0	788	14	TMO308685	AJ308685 Tobacco m
8	76	100.0	790	14	TMO307583	AJ307583 Tobacco m
9	76	100.0	804	14	TMO307582	AJ307582 Tobacco m
10	76	100.0	806	14	TMO308682	AJ308682 Tobacco m
11	76	100.0	806	14	TMO308684	AJ308684 Tobacco m
12	76	100.0	807	6	BD263879	BD263879 Viral exp
13	76	100.0	807	6	BD263880	BD263880 Viral exp
14	76	100.0	807	6	AR435699	AR435699 Sequence
15	76	100.0	807	6	AR435700	AR435700 Sequence
16	76	100.0	807	6	AX045758	AX045758 Sequence
17	76	100.0	807	6	AX592974	AX592974 Sequence
18	76	100.0	807	14	TMO307581	AJ307581 Tobacco m
19	76	100.0	807	14	TMO308683	AJ308683 Tobacco m

20	76	100.0	807	14	TMO308693	AJ308693 Tobacco m
21	76	100.0	807	14	TMO309081	AJ509081 Tobacco m
22	76	100.0	808	14	TMO307578	AJ307578 Tobacco m
23	76	100.0	808	14	TMO308686	AJ308686 Tobacco m
24	76	100.0	808	14	TMO308688	AJ308688 Tobacco m
25	76	100.0	808	14	TMO308690	AJ308690 Tobacco m
26	76	100.0	809	14	TMO308691	AJ308691 Tobacco m
27	76	100.0	811	14	TMO309082	AJ509082 Tobacco m
28	76	100.0	817	14	TMO309083	AJ509083 Tobacco m
29	76	100.0	857	6	AX045757	AX045757 Sequence
30	76	100.0	891	14	TMO307580	AJ307580 Tobacco m
31	76	100.0	891	14	TMO308687	AJ308687 Tobacco m
32	76	100.0	1825	6	AR042908	AR042908 Sequence
33	76	100.0	1825	6	AR079857	AR079857 Sequence
34	76	100.0	6395	6	AR173320	AR173320 Sequence
35	76	100.0	6395	6	AR271575	AR271575 Sequence
36	76	100.0	6395	6	AX040174	AX040174 Sequence
37	76	100.0	6395	6	AX098414	AX098414 Sequence
38	76	100.0	6395	14	TOTMV4	V01408 Tobacco mos
39	76	100.0	6398	14	TOTMV5	V01409 Tobacco mos
40	76	100.0	6425	6	AR173322	AR173322 Sequence
41	76	100.0	6425	6	AX098416	AX098416 Sequence
42	76	100.0	6439	6	AR173321	AR173321 Sequence
43	76	100.0	6439	6	AX098415	AX098415 Sequence
44	76	100.0	6446	6	AR173324	AR173324 Sequence
45	76	100.0	6446	6	AX098418	AX098418 Sequence

ALIGNMENTS

RESULT 1
TMO307579
LOCUS
DEFINITION
Tobacco mosaic virus MP gene for movement protein, genomic RNA, lethal mutant A23.
ACCESSION
AJ307579
VERSION
AJ307579.1 GI:15551731
KEYWORDS
movement protein; mp gene.
SOURCE
Tobacco mosaic virus
ORGANISM
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE
1
AUTHORS
Malpica,J.M., Fraile,A., Obies,C.I., Drake,J.W. and Garcia-Arenal,F.
TITLE
The rate and character of spontaneous mutation in an RNA virus
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 396)
AUTHORS
Fraile,A.
TITLE
Direct Submission
JOURNAL
Submitted (08-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN
FEATURES
location/Qualifiers
1..396
source
organism="Tobacco mosaic virus"
/virion
/mol_type="genomic RNA"
/db_xref="taxon:12242"
/note="lethal mutant A23"
1..396
gene
/gene="Mp"
/pseudo
1..396
CDS
/gene="Mp"
/pseudo
/codon_start=1
/product="movement protein"
/db_xref="PSEUDO:CAC69430.1"

ORIGIN

Query Match 100.0%; Score 76; DB 14; Length 396;
Best Local Similarity 100.0%; Pred. No. 3.5e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
 |||||
 Db 130 GTGAGACGGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 189
 |||||
 Qy 61 GTCCTATGTCGATCA 76
 |||||
 Db 190 GTCCTATGTCGATCA 205
 |||||

RESULT 2
 A68760
 LOCUS A68760 632 bp DNA linear PAT 06-MAY-1999
 DEFINITION Sequence 1 from Patent WO9803668.
 ACCESSION A68760
 VERSION A68760.1 GI:4759734
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified

REFERENCE 1 (bases 1 to 632)
 AUTHORS Atabekov,I.G., Dorokhov,Y.L. and Morozov,S.Y.
 TITLE VIRUS RESISTANCE IN PLANTS
 JOURNAL Patent: WO 9803668-A 1 29-JAN-1998;
 ZENECA LTD (GB)
 FEATURES Location/Qualifiers
 source 1..632
 /organism="unidentified"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32644"
 /clone="TMV COAT PROTEIN GENE SUB-GENOMIC PROMOTER"

ORIGIN
 Query Match 100.0%; Score 76; DB 6; Length 632;
 Best Local Similarity 100.0%; Pred. No. 3.6e-13;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
 |||||
 Db 364 GTGAGACGGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 423
 |||||

Qy 61 GTCCTATGTCGATCA 76
 |||||
 Db 424 GTCCTATGTCGATCA 439
 |||||

RESULT 3
 TMO308689
 LOCUS TMO308689 706 bp RNA linear VRL 08-SEP-2001
 DEFINITION Tobacco mosaic virus MP gene for movement protein, genomic RNA, lethal mutant B1192.

ACCESSION AJ308689
 VERSION AJ308689.1 GI:15552815
 KEYWORDS movement protein; mp gene.
 SOURCE Tobacco mosaic virus
 ORGANISM Tobacco mosaic virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.

REFERENCE 1
 AUTHORS Malpica,J.M., Fraile,A., Obies,C.I., Drake,J.W. and Garcia-Arenal,F.
 TITLE The rate and character of spontaneous mutation in an RNA virus

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 706)
 AUTHORS Fraile,A.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN

FEATURES Location/Qualifiers
 source 1..706
 /organism="Tobacco mosaic virus"
 /viral
 /mol_type="genomic RNA"
 /db_xref="taxon:12242"

/notes="lethal mutant B1192"

1..706
 /gene="MP"
 /pseudo
 1..706
 /gene="MP"
 /pseudo
 /codon_start=1
 /product="movement protein"

ORIGIN

Query Match 100.0%; Score 76; DB 14; Length 706;
 Best Local Similarity 100.0%; Pred. No. 3.6e-13;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
 |||||
 Db 440 GTGAGACGGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 499
 |||||

Qy 61 GTCCTATGTCGATCA 76
 |||||
 Db 500 GTCCTATGTCGATCA 515
 |||||

RESULT 4

TMO509080
 LOCUS TMO509080 750 bp RNA linear VRL 30-OCT-2002
 DEFINITION Tobacco mosaic virus MP pseudogene, genomic RNA, lethal mutant B216.

ACCESSION AJ509080
 VERSION AJ509080.1 GI:24459913
 KEYWORDS movement protein; MP, pseudo gene.
 SOURCE Tobacco mosaic virus
 ORGANISM Tobacco mosaic virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.

REFERENCE 1
 AUTHORS Malpica,J.M., Fraile,A., Moreno,I.M., Obies,C.I., Drake,J.M. and Garcia-Arenal,F.
 TITLE The rate and character of spontaneous mutation in an RNA virus
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 750)
 AUTHORS Fraile,A.

Direct Submission
 TITLE Submitted (27-SEP-2002) Fraile A., Biotechnology, E.T.S.Ingenieros Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN

FEATURES Location/Qualifiers
 source 1..750
 /organism="Tobacco mosaic virus"
 /viral
 /mol_type="genomic RNA"
 /isolate="lethal mutant B216"
 /db_xref="taxon:12242"

1..750
 /gene="MP"
 /pseudo
 1..750
 /gene="MP"
 /pseudo
 /codon_start=1
 /db_xref="PSEUDO:CAD48853.1"

ORIGIN

Query Match 100.0%; Score 76; DB 14; Length 750;
 Best Local Similarity 100.0%; Pred. No. 3.6e-13;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
 |||||
 Db 484 GTGAGACGGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 543
 |||||

Qy 61 GTCCTATGTCGATCA 76
 |||||
 Db 544 GTCCTATGTCGATCA 559
 |||||

```
RESULT 5
TMO309084
LOCUS       TMO309084            750 bp    RNA        linear    VRL 30-OCT-2002
DEFINITION  Tobacco mosaic virus MP pseudogene, genomic RNA, lethal mutant B8.
ACCESSION   AJ509084
VERSION     AJ509084.1 GI:24459918
KEYWORDS    movement protein; MP, pseudo gene.
SOURCE      Tobacco mosaic virus
ORGANISM    Tobacco mosaic virus
REFERENCE   1
AUTHORS     Malpica,J.M.; Fraile,A., Moreno,I.M., Obies,C.I., Drake,J.W. and
            Garcia-Arenal,F.
TITLE       The rate and character of spontaneous mutation in an RNA virus
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 750)
AUTHORS     Fraile,A.
TITLE       Direct Submission
JOURNAL     Submitted (27-SEP-2002) Fraile A., Biotechnology, E.T.S.Ingenieros
            Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN
FEATURES    Location/Qualifiers
             source          1..750
             /organism="Tobacco mosaic virus"
             /mol_type="genomic RNA"
             /isolate="lethal mutant B8"
             /db_xref="taxon:12242"
             gene            1..750
             /gene="Mp"
             /pseudo
             CDS              1..750
             /gene="Mp"
             /pseudo
             /codon_start=1
             /db_xref="PSEUDO:CAD48857.1"
ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 750;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 60
Db 484 GTGAGACGCGAGGCCCATCGAACTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 543

Qy 61 GTCCTATGTCGATCA 76
Db 544 GTCCTATGTCGATCA 559

RESULT 6
TMO308692
LOCUS       TMO308692            782 bp    RNA        linear    VRL 08-SEP-2001
DEFINITION  Tobacco mosaic virus MP gene for movement protein, genomic RNA,
            lethal mutant B303.
ACCESSION   AJ308692
VERSION     AJ308692.1 GI:15552818
KEYWORDS    movement protein; mp gene.
SOURCE      Tobacco mosaic virus
ORGANISM    Tobacco mosaic virus
REFERENCE   1
AUTHORS     Malpica,J.M.; Fraile,A., Obies,C.I., Drake,J.W. and
            Garcia-Arenal,F.
TITLE       The rate and character of spontaneous mutation in an RNA virus
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 782)
AUTHORS     Fraile,A.
TITLE       Direct Submission
JOURNAL     Submitted (09-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros
            Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN
FEATURES    Location/Qualifiers
             source          1..782
             /organism="Tobacco mosaic virus"
             /mol_type="genomic RNA"
             /db_xref="taxon:12242"
             /note="lethal mutant B1084"
             gene            1..788
             /gene="Mp"
             /pseudo
             CDS              1..788
             /gene="Mp"
             /pseudo
             /codon_start=1
             /product="movement protein"
ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 788;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 60
```

```
FEATURES    Location/Qualifiers
             source          1..782
             /organism="Tobacco mosaic virus"
             /mol_type="genomic RNA"
             /db_xref="taxon:12242"
             /note="lethal mutant B303"
             gene            1..782
             /gene="Mp"
             /pseudo
             CDS              1..782
             /gene="Mp"
             /pseudo
             /codon_start=1
             /product="movement protein"
ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 782;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 60
Db 540 GTGAGACGCGAGGCCCATCGAACTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 599

Qy 61 GTCCTATGTCGATCA 76
Db 600 GTCCTATGTCGATCA 615

RESULT 7
TMO308685
LOCUS       TMO308685            788 bp    RNA        linear    VRL 08-SEP-2001
DEFINITION  Tobacco mosaic virus MP gene for movement protein, genomic RNA,
            lethal mutant B1084.
ACCESSION   AJ308685
VERSION     AJ308685.1 GI:15552811
KEYWORDS    movement protein; mp gene.
SOURCE      Tobacco mosaic virus
ORGANISM    Tobacco mosaic virus
REFERENCE   1
AUTHORS     Malpica,J.M.; Fraile,A., Obies,C.I., Drake,J.W. and
            Garcia-Arenal,F.
TITLE       The rate and character of spontaneous mutation in an RNA virus
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 788)
AUTHORS     Fraile,A.
TITLE       Direct Submission
JOURNAL     Submitted (09-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros
            Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN
FEATURES    Location/Qualifiers
             source          1..788
             /organism="Tobacco mosaic virus"
             /mol_type="genomic RNA"
             /db_xref="taxon:12242"
             /note="lethal mutant B1084"
             gene            1..788
             /gene="Mp"
             /pseudo
             CDS              1..788
             /gene="Mp"
             /pseudo
             /codon_start=1
             /product="movement protein"
ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 788;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 60
```

```
|||||
522 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 581
Qy 61 GTCCTATGTCGATCA 76
Db 582 GTCCTATGTCGATCA 597

RESULT 8
TWO307583
LOCUS      790 bp      RNA      linear      VRL 08-SEP-2001
DEFINITION Tobacco mosaic virus MP gene for movement protein, genomic RNA,
            lethal mutant B895.
ACCESSION  AJ307583
VERSION     AJ307583.1 GI:15552807
KEYWORDS   movement protein; mp gene.
SOURCE     Tobacco mosaic virus
ORGANISM   Tobacco mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE  1
AUTHORS    Malpica,J.M., Fraile,A., Obies,C.I., Drake,J.W. and
            Garcia-Arenal,F.
TITLE      The rate and character of spontaneous mutation in an RNA virus
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 790)
AUTHORS    Fraile,A.
DIRECT SUBMISSION
TITLE      Submitted (08-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros
JOURNAL    Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN
FEATURES   source
            1..790
            /organism="Tobacco mosaic virus"
            /viroion
            /mol_type="genomic RNA"
            /db_xref="taxon:12242"
            /note="lethal mutant B895"
            1..790
            /gene="MP"
            /pseudo
            1..790
            /gene="MP"
            /pseudo
            /codon_start=1
            /product="movement protein"

ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 790;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 60
    |||||||
Db 523 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 582
    |||||||

Qy 61 GTCCTATGTCGATCA 76
Db 583 GTCCTATGTCGATCA 598

RESULT 9
TWO307582
LOCUS      804 bp      RNA      linear      VRL 08-SEP-2001
DEFINITION Tobacco mosaic virus MP gene for movement protein, genomic RNA,
            lethal mutant B420.
ACCESSION  AJ307582
VERSION     AJ307582.1 GI:15552806
KEYWORDS   movement protein; mp gene.
SOURCE     Tobacco mosaic virus
ORGANISM   Tobacco mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE  1
AUTHORS    Malpica,J.M., Fraile,A., Obies,C.I., Drake,J.W. and
            Garcia-Arenal,F.
TITLE      The rate and character of spontaneous mutation in an RNA virus
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 806)
AUTHORS    Fraile,A.
DIRECT SUBMISSION
TITLE      Submitted (09-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros
JOURNAL    Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN
FEATURES   source
            1..806
            /organism="Tobacco mosaic virus"
            /viroion
            /mol_type="genomic RNA"
            /db_xref="taxon:12242"
            /note="lethal mutant A292"
            1..806
            /gene="MP"
            /pseudo
            1..806
            /gene="MP"
            /pseudo
            /codon_start=1

ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 790;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 60
    |||||||
Db 523 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 582
    |||||||

Qy 61 GTCCTATGTCGATCA 76
Db 583 GTCCTATGTCGATCA 598
```

```
The rate and character of spontaneous mutation in an RNA virus
Unpublished
REFERENCE  2 (bases 1 to 804)
AUTHORS    Fraile,A.
DIRECT SUBMISSION
TITLE      Submitted (08-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros
JOURNAL    Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN
FEATURES   Location/Qualifiers
            source
            1..804
            /organism="Tobacco mosaic virus"
            /viroion
            /mol_type="genomic RNA"
            /db_xref="taxon:12242"
            /note="lethal mutant B420"
            1..804
            /gene="MP"
            /pseudo
            1..804
            /gene="MP"
            /pseudo
            /codon_start=1
            /product="movement protein"
            /db_xref="PSEUDO:CAC69433.1"

ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 804;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 60
    |||||||
Db 538 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 597
    |||||||

Qy 61 GTCCTATGTCGATCA 76
Db 598 GTCCTATGTCGATCA 613

RESULT 10
TWO308682
LOCUS      806 bp      RNA      linear      VRL 08-SEP-2001
DEFINITION Tobacco mosaic virus MP gene for movement protein, genomic RNA,
            lethal mutant A292.
ACCESSION  AJ308682
VERSION     AJ308682.1 GI:15552808
KEYWORDS   movement protein; mp gene.
SOURCE     Tobacco mosaic virus
ORGANISM   Tobacco mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE  1
AUTHORS    Malpica,J.M., Fraile,A., Obies,C.I., Drake,J.W. and
            Garcia-Arenal,F.
TITLE      The rate and character of spontaneous mutation in an RNA virus
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 806)
AUTHORS    Fraile,A.
DIRECT SUBMISSION
TITLE      Submitted (09-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros
JOURNAL    Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN
FEATURES   Location/Qualifiers
            source
            1..806
            /organism="Tobacco mosaic virus"
            /viroion
            /mol_type="genomic RNA"
            /db_xref="taxon:12242"
            /note="lethal mutant A292"
            1..806
            /gene="MP"
            /pseudo
            1..806
            /gene="MP"
            /pseudo
            /codon_start=1

ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 790;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 60
    |||||||
Db 523 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGTCGTTGATGAGTTCATGGAAGAT 582
    |||||||

Qy 61 GTCCTATGTCGATCA 76
Db 583 GTCCTATGTCGATCA 598
```



```
ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 806;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db 540 GTGAGACGCGAGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 599

Qy 61 GTCCTATGTCGATCA 76
Db 600 GTCCTATGTCGATCA 615

RESULT 11
TMO308684
LOCUS      806 bp      RNA      linear      VRL 08-SEP-2001
DEFINITION Tobacco mosaic virus MP gene for movement protein, genomic RNA,
lethal mutant A939.
ACCESSION  AJ308684
VERSION     AJ308684.1 GI:15552810
KEYWORDS   movement protein; mp gene.
SOURCE     Tobacco mosaic virus
ORGANISM   Tobacco mosaic virus
REFERENCE  1
AUTHORS    Malpica,J.M., Fraile,A., Obies,C.I., Drake,J.W. and
Garcia-Arenal,F.
TITLE      The rate and character of spontaneous mutation in an RNA virus
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 806)
AUTHORS    Fraile,A.
TITLE      Direct Submission
JOURNAL    Submitted (09-MAR-2001) Fraile A., Biotechnology, E.T.S.Ingenieros
Agronomos, Ciudad Universitaria, 28040 Madrid, SPAIN

FEATURES
source
1..806
/organism="Tobacco mosaic virus"
/virus
/mol_type="genomic RNA"
/db_xref="taxon:12242"
/note="lethal mutant A939"
1..806
/gene="mp"
/pseudo
1..806
/gene="mp"
/pseudo
/codon_start=1
/product="movement protein"

ORIGIN
Query Match      100.0%; Score 76; DB 14; Length 806;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db 540 GTGAGACGCGAGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 599

Qy 61 GTCCTATGTCGATCA 76
Db 600 GTCCTATGTCGATCA 615

RESULT 12
BD263879
LOCUS      807 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Viral expression vectors.
ACCESSION  BD263879
VERSION     BD263879.1 GI:33073647
```

```
KEYWORDS
SOURCE      JP 2002542828-A/3.
ORGANISM    Nicotiana tabacum (common tobacco)

REFERENCE
AUTHORS     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE       Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
JOURNAL     asterids; lamids; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 807)
Fitzmaurice,W.P., Pogue,G.P. and Lindbo,J.A.
Patent: JP 2002542828-A 3 17-DEC-2002;
LARGE SCALE BIOLOGY CORP
OS          Nicotiana tabacum (tobacco)
PN          JP 2002542828-A/3
PD          17-DEC-2002
PF          04-MAY-2000 JP 2000615766
PI          04-MAY-1999 US 60/132697
PT          WAYNE P FITZMAURICE,GREGORY P POGUE,JOHN A LINDBO PC
C12N15/09,A01H5/00,C12N5/10,C12N5/00,C12N5/00 CC Viral
expression vectors
FH Key      Location/Qualifiers
FT source   1..807
            /organism="Nicotiana tabacum (tobacco)"

FEATURES
source
1..807
/organism="Nicotiana tabacum"
/mol_type="genomic DNA"
/db_xref="taxon:4097"

ORIGIN
Query Match      100.0%; Score 76; DB 6; Length 807;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db 541 GTGAGACGCGAGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 13
BD263880
LOCUS      807 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Viral expression vectors.
ACCESSION  BD263880
VERSION     BD263880.1 GI:33073648
KEYWORDS   JP 2002542828-A/4.
SOURCE     Nicotiana tabacum (common tobacco)
ORGANISM    Nicotiana tabacum

REFERENCE
AUTHORS     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE       Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
JOURNAL     asterids; lamids; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 807)
Fitzmaurice,W.P., Pogue,G.P. and Lindbo,J.A.
Patent: JP 2002542828-A 4 17-DEC-2002;
LARGE SCALE BIOLOGY CORP
OS          Nicotiana tabacum (tobacco)
PN          JP 2002542828-A/4
PD          17-DEC-2002
PF          04-MAY-2000 JP 2000615766
PI          04-MAY-1999 US 60/132697
PT          WAYNE P FITZMAURICE,GREGORY P POGUE,JOHN A LINDBO PC
C12N15/09,A01H5/00,C12N5/10,C12N5/00,C12N5/00 CC Viral
expression vectors
FH Key      Location/Qualifiers
FT source   1..807
            /organism="Nicotiana tabacum (tobacco)"

FEATURES
source
1..807
/organism="Nicotiana tabacum"
/mol_type="genomic DNA"
/db_xref="taxon:4097"
```

```
/mol_type="genomic DNA"  
/db_xref="taxon:4097"
```

```

ORIGIN
Query Match          100.0%;   Score 76;   DB 6;   Length 807;
Best Local Similarity 100.0%;   Pred. No. 3.6e-13;
Matches 76;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy  1  GTGAGACGCGAGGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT  60
      |||||
Db   541 GTGAGACGCGAGGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT  600
      |||||

Qy  61  GTCCTATGTCGATCA  76
      |||||
Db   601 GTCCTATGTCGATCA  616
      |||||

```

```

RESULT 14
AR435699
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
AR435699
Sequence 3 from patent US 6656726.
AR435699
AR435699.1 GI:40198773
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 807)
Fitzmaurice,W.P., Pogue,G.P. and Lindbo,J.A.
Viral expression vectors
Patent: US 6656726-A 3 02-DEC-2003;
Location/Qualifiers
1..807
/organism="unknown"
/mol_type="genomic DNA"

```

Query Match	100.0%	Score 76;	DB 6;	Length 807;
Best Local Similarity	100.0%;	Pred. No. 3.6e-13;		
Matches	76;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	GTGAGACGGAGGGCCCATGGAACTTACAGAGAAGTCGTTGATGAGTTTCATGGAGAT	60	
Db	541	GTGAGACGGAGGGCCCATGGAACTTACAGAGAAGTCGTTGATGAGTTTCATGGAGAT	600	
Qy	61	GTCCCTATGTCGATCA	76	
Db	601	GTCCCTATGTCGATCA	616	

RESULT_15		PAT 18-DEC-2003
AR435700	Sequence 4 from patent US 6656726.	DNA linear
Locus	AR435700	807 bp
DEFINITION	Accession AR435700	
VERSION	AR435700.1 GI:40198774	
KEYWORDS	.	
SOURCE	. Unknown.	
ORGANISM	Unknown.	
REFERENCE	Unclassified.	
AUTHORS	1 (bases 1 to 807)	
TITLE	Fitzmaurice,W.P., Fogue,G,P. and Lindbo,J.A.	
JOURNAL	viral expression vectors	
FEATURES	Patent: US 6656726-A 4 02-DEC-2003;	
	Location/Qualifiers	
	1..807	
	/organism='unknown'	'DNA'
	/mol_type='genomic'	
	source	

Query Match	100.0%	Score 76;	DB 6;	Length 807;
Best Local Similarity	100.0%	Pred. No. 3.6e-13;		
Matches 76;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Search completed: September 26, 2005, 03:05:05
Job time : 535.111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 23:45:48 ; Search time 95.5556 Seconds
(without alignments)
4708.259 Million cell updates/sec

Title: US-09-551-494-2_COPY_5443_5518

Perfect score: 76

Sequence: 1 gtgagagacggaggcccat.....agatgtccctatgtcgatca 76

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : N_Geneseq_16Dec04:*
- 1: Geneseqn1980s:*
 - 2: Geneseqn1990s:*
 - 3: Geneseqn2000s:*
 - 4: Geneseqn2001as:*
 - 5: Geneseqn2001bs:*
 - 6: Geneseqn2002as:*
 - 7: Geneseqn2002bs:*
 - 8: Geneseqn2003as:*
 - 9: Geneseqn2003bs:*
 - 10: Geneseqn2003cs:*
 - 11: Geneseqn2003ds:*
 - 12: Geneseqn2004as:*
 - 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	76	100.0	557	ADD17931	Add17931 DNA (Seqi
c 2	76	100.0	557	ADK56952	Adk56952 Plant DNA
3	76	100.0	632	AAV16847	Aav16847 Tobacco m
4	76	100.0	807	ABT06572	Abt06572 Wild-type
5	76	100.0	807	ABX14577	Abx14577 Tomato mo
6	76	100.0	807	ADM68440	Adm68440 Tobacco m
7	76	100.0	807	ADP26593	Adp26593 Tobamovir
8	76	100.0	807	ADQ88415	Adq88415 TMV-U1 mo
9	76	100.0	807	ADS17169	Ads17169 Tobacco m
10	76	100.0	895	AAQ62653	Aaq62653 Tobacco m
11	76	100.0	1187	ADD17932	Add17932 DNA (Seqi
12	76	100.0	1187	ADK56953	Adk56953 Plant DNA
13	76	100.0	1825	AAQ67663	Aaq67663 TMV repli
14	76	100.0	1825	AAV54825	Aav54825 Replicon
15	76	100.0	5484	ADA15011	Ada15011 Tobacco m
16	76	100.0	6395	2 AAQ95155	Aaq95155 Tobacco m
17	76	100.0	6395	2 AAZ20642	Aaz20642 TMV-based
18	76	100.0	6395	3 AAC62369	Aac62369 cDNA sequ
19	76	100.0	6395	4 AAF82330	Aaf82330 Wild-type
20	76	100.0	6425	2 AAZ20644	Aaz20644 TMV-based

21	76	100.0	6425	4 AAF82332	Aaf82332 Tobacco m
22	76	100.0	6439	2 AAZ20643	Aaz20643 TMV-based
23	76	100.0	6439	4 AAF82331	Aaf82331 Tobacco m
24	76	100.0	6446	2 AAZ20646	Aaz20646 TMV-based
25	76	100.0	6446	4 AAF82334	Aaf82334 Tobacco m
26	76	100.0	6475	2 AAZ20645	Aaz20645 TMV-based
27	76	100.0	6475	4 AAF82333	Aaf82333 Tobacco m
28	76	100.0	7684	8 ADA15014	Ada15014 Tobacco m
29	76	100.0	7684	10 ADB83322	Adb83322 DNA trans
30	76	100.0	7685	4 AAD02010	Aad02010 TMV viral
31	76	100.0	7685	5 AAD20298	Aad20298 Tobacco m
32	76	100.0	7685	5 AAD20295	Aad20295 Tobacco m
33	76	100.0	7685	5 AAD20288	Aad20288 Tobacco m
34	76	100.0	7685	5 AAD20291	Aad20291 Tobacco m
35	76	100.0	7685	5 AAD20296	Aad20296 Tobacco m
36	76	100.0	7685	5 AAD20297	Aad20297 Tobacco m
37	76	100.0	7685	5 AAD20294	Aad20294 Tobacco m
38	76	100.0	7685	6 AAD24472	Aad24472 Tobacco m
39	76	100.0	7685	6 AAD24480	Aad24480 Tobacco m
40	76	100.0	7685	6 AAD24479	Aad24479 Tobacco m
41	76	100.0	7685	6 AAD24478	Aad24478 Tobacco m
42	76	100.0	7685	6 AAD24481	Aad24481 Tobacco m
43	76	100.0	7685	6 AAD24475	Aad24475 Tobacco m
44	76	100.0	7685	6 AAD24482	Aad24482 Tobacco m
45	76	100.0	7685	8 ADA15015	Ada15015 Tobacco m

ALIGNMENTS

RESULT 1

ADD17931/c

ID ADD17931 standard; DNA; 557 BP.

XX AC ADD17931;

XX DT 15-JAN-2004 (first entry)

XX DE DNA (SeqID 1999) that confers an altered visual phenotype in plants.

XX KW ds; visual phenotype; plant; architecture; leaf surface; chlorotic;

XX KW bleaching; etching; wet leaf; stunting; elongation; texture;

XX KW agronomic trait; growth regulation; dwarf variety; insect resistance;

XX KW heat stress; transgenic.

OS Unidentified.

XX FN W02003020741-Al..

XX PD 13-MAR-2003.

XX PF 30-AUG-2002; 2002WO-US027880.

XX PR 31-AUG-2001; 2001US-0316326P.

XX PA (DOWC) DOW CHEM CO.

XX PA (DOWC) DOW AGROSCIENCES LLC.

XX PI Crosley R, Skokut T, Ruegger M, Larrinua I, Shukla V;

XX DR WPI; 2003-300858/29.

XX PT Novel isolated nucleic acid derived from Nicotiana benthamiana. Oryzae

XX PT sativa, Saccharomyces cerevisiae, and Papaver rhoeas, useful for

XX PT conferring altered visual phenotypes in plants.

XX PS Claim 1; SEQ ID NO 1999; 517pp; English.

XX CC This invention relates to the identification and isolation of novel
XX CC nucleic acid molecules that confer altered visual phenotypes in plants.
XX CC Specifically, it refers to modifications of plant architecture and/or
XX CC leaf surface features in plants, such as chlorotic, bleaching, etching,
XX CC wet leaf, stunting, elongation and texture phenotypes, which are thought

CC will be agronomic traits beneficial to the farmer. As such, these novel
CC phenotypes can affect growth regulation i.e. useful for creating dwarf
CC varieties, exhibit resistance to insects or heat stress, confer changes
CC in pigment content such that plants have enhanced vitamin production or
CC delayed senescence and also for example produce plants that control the
CC production of ethylene. Furthermore, the present invention comprises
CC generating transgenic plants, as well as reproducibly altering the visual
CC phenotype of plant seeds, plant tissues and plant cells containing the
CC polynucleotides described herein. This polynucleotide is a homologue of a
CC DNA sequence that confers an altered visual phenotype when expressed in
CC plants, the method of the invention.

XX Sequence 557 BP; 155 A; 143 C; 87 G; 172 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 10; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAGACGCGGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
DB 373 GTGAGACGCGGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 314
QY 61 GTCCTATGTCGATCA 76
DB 313 GTCCTATGTCGATCA 298

RESULT 2
ADK56952/c
ID ADK56952 standard; DNA; 557 BP.
XX
AC ADK56952;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plant DNA sequence which confers altered metabolic characteristic #4335.
XX
KW altered metabolic characteristic; plant; acid metabolism;
KW alcohol metabolism; fatty acid metabolism;
KW branched fatty acid metabolism; alkaloid metabolism;
KW amino acid metabolism; ester metabolism; glyceride metabolism;
KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
OS Unidentified.
XX
XX WO2003020936-A1.
XX
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027884.
XX
XX 31-AUG-2001; 2001US-0316471P.
XX
XX (DOWC) DOW CHEM CO.
XX (DOWC) DOW AGROSCIENCES LLC.
XX
XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
XX Oriedo JVB, Croasley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX WPI; 2003-313091/30.
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
XX benhamiana plants, useful for altering the levels of metabolites e.g.
XX acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX Claim 1; SEQ ID NO 4335; 2576pp; English.

XX The invention comprises DNA sequences which confer an altered metabolic
XX characteristic when they are expressed in a plant. The DNA sequences of
XX the invention are useful for producing plants with an altered metabolic

CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC altered glyceride metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.

SQ Sequence 557 BP; 155 A; 143 C; 87 G; 172 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 10; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAGACGCGGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
DB 373 GTGAGACGCGGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 314
QY 61 GTCCTATGTCGATCA 76
DB 313 GTCCTATGTCGATCA 298

RESULT 3
AAV16847
ID AAV16847 standard; DNA; 632 BP.
XX
AC AAV16847;
XX
DT 24-JUN-1998 (first entry)
XX
DE Tobacco mosaic virus coat protein gene sub-genomic promoter.
XX
KW TMV; coat protein gene; sub-genomic promoter; plant virus; inhibition;
KW replication; resistance; virus-specific RNA dependent RNA polymerase;
XX RDRP; replication; ss.
XX
OS Tobacco mosaic virus.
XX
XX WO9803668-A1.
XX
XX 29-JAN-1998.
XX
XX 21-JUL-1997; 97WO-GB001960.
XX
XX 22-JUL-1996; 96GB-00015349.
XX
XX (ZENE) ZENECA LTD.
XX
XX Atabekov IG, Dorokhov YL, Morozov SY;
XX WPI; 1998-120787/11.
XX
XX DNA constructs improving or imparting plant resistance to viruses - by
XX inhibiting virus replication, useful for producing transgenic plants e.g.
XX tobacco with increased inhibition of viral replication.
XX
XX Claim 2; Page 15; 35pp; English.

XX The present sequence represents the Tobacco mosaic virus (TMV) coat
XX protein gene sub-genomic promoter. The specification describes a new DNA
XX construct (which is specifically claimed) which comprises a promoter
XX operable in plants, a polynucleotide with a sequence encoding an RNA
XX which is capable of binding to a plant virus RNA dependent polymerase,
XX and which is the promoter region of a gene that is subgenomic when in the
XX viral genome (e.g. present sequence), and a terminator sequence
XX heterologous to the promoter. A method for imparting and improving the
XX ability of plants to inhibit replication of infecting viruses is
XX described. The present sequence is preferably repeated at least once, and
XX repeats arranged in tandem in the construct. Plants with an improved

CC ability to inhibit the replication of infecting viruses can be produced
 CC using the constructs. Such plants may have improved resistance to either
 CC one or several viruses. The construct allows production, in transgenic
 CC plants, of an RNA transcript mimicking binding sites of the virus-
 CC specific RNA dependent RNA polymerase (RDRP) which is produced by the
 CC infecting virus. The transgenic binding sites compete for the RDRP, and
 CC binding of this to the transgenic sequence rather than the infecting
 CC virus results in the inability of the infecting virus to successfully
 CC replicate

XX SQ Sequence 632 BP; 193 A; 98 C; 172 G; 169 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 2; Length 632;
 Best Local Similarity 100.0%; Pred. No. 2.6e-16;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 60
 |||||
 DB 364 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 423
 |||||
 OY 61 GTCCTATGTCGATCA 76
 |||||
 DB 424 GTCCTATGTCGATCA 439
 |||||

RESULT 4
 ABT06572
 ID ABT06572 standard; DNA; 807 BP.

XX AC ABT06572;

XX DT 30-OCT-2002 (first entry)

XX DE Wild-type viral MP gene sequence SEQ ID No 17.

XX KW Recombination; double-stranded product; homology-driven reassembly;
 KW parental template; gene therapy; replication-defective; ds.

XX OS Tobamovirus.

XX PN WO200261142-A2.

XX PD 08-AUG-2002.

XX PF 31-JAN-2002; 2002WO-US003089.

XX PR 31-JAN-2001; 2001US-00775049.

XX PA (LARG-) LARGE SCALE BIOLOGY CORP.

XX PI Padgett HS, Fitzmaurice WP, Lindbo JA;

XX DR WPI; 2002-627485/67.

XX Forcing recombination between polynucleotides, useful for e.g. shuffling
 PT viral genes or sequences to generate gene therapy vectors, by using a
 PT single strand of parental template for the homology-driven reassembly of
 PT nucleic acids.

XX PS Example 2; Page 75; 132pp; English.

XX The invention relates to methods for forcing recombination between
 CC polynucleotides. This involves employing a single strand of each input
 CC parental template for repeated cycles of reassembly, recombination and
 CC selection in order for double-stranded products to result. The methods
 CC are useful for the homology-driven reassembly of nucleic acid sequences.
 CC In particular, the method is useful for forcing recombination among input
 CC parental templates such that reassembly produces randomised
 CC polynucleotides that can be screened for a particular use. The method is
 CC therefore useful for generating or producing large libraries of widely
 CC varying mutant nucleic acid sequences, typically in the form of amplified
 CC and/or cloned polynucleotides. The selected polynucleotide sequence(s)
 CC possesses at least one desired phenotypic characteristic (e.g. encodes a

CC polypeptide, promotes transcription of linked polynucleotides, binds a
 CC protein, or improves the function of a viral vector), which can be
 CC selected or screened for. The method may also be used for e.g. shuffling
 CC a population of viral genes, shuffling polynucleotide sequence for
 CC generating gene therapy vectors or replication-defective gene therapy
 CC constructs, or to test for optimised characteristics of a nucleic acid or
 CC polypeptide. This polynucleotide sequence represents a wild-type viral MP
 CC gene sequence relating to the recombination methods of the invention

XX SQ Sequence 807 BP; 251 A; 121 C; 213 G; 222 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 6; Length 807;
 Best Local Similarity 100.0%; Pred. No. 2.8e-16;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 60
 |||||
 DB 541 GTGAGACGCGAGGCCCATGGAACCTTACAGAAGAAGTCGTTGATGAGTTCATGGAAGAT 600
 |||||
 OY 61 GTCCTATGTCGATCA 76
 |||||
 DB 601 GTCCTATGTCGATCA 616
 |||||

RESULT 5
 ABX14577
 ID ABX14577 standard; DNA; 807 BP.

XX AC ABX14577;

XX DT 14-MAR-2003 (first entry)

XX DE Tomato mosaic virus (ToMV) movement protein gene.

XX KW Tomato mosaic virus; gene; ds; heteroduplex; T4 DNA polymerase; ToMV;
 KW T4 DNA ligase; CEL I; population diversity; exonuclease; complementarity;
 KW movement protein.

XX OS Tomato mosaic virus.

XX PN WO200279468-A2.

XX PD 10-OCT-2002.

XX PF 01-FEB-2002; 2002WO-US003055.

XX PR 02-FEB-2001; 2001US-0266386P.

XX PR 14-FEB-2001; 2001US-0268785P.

XX PA (LARG-) LARGE SCALE BIOLOGY CORP.

XX PI Padgett HS, Fitzmaurice WP, Lindo JA;

XX DR WPI; 2003-092898/08.

XX Increasing complementarity by making sequence variants from heteroduplex
 PT polynucleotides, useful for producing improved gene products from
 PT randomly mutated genes.

XX PS Example 10; Page 102-103; 105pp; English.

XX The invention relates to an in vitro method of making sequence variants
 CC from at least one heteroduplex polynucleotide that has at least two-non
 CC complementary nucleotide base pairs. The method comprises combining a
 CC heteroduplex polynucleotide with CEL I, T4 DNA polymerase and T4 DNA
 CC ligase, or an agent or agents with exonuclease activity and allowing
 CC sufficient time for the percentage of complementarity to increase, where
 CC one or more variants are made. The method can also be used to increase
 CC diversity in a population of sequences. The method is useful for
 CC producing improved gene products from randomly mutated genes or from
 CC expression from a suitable plant, animal, fungal, yeast or bacterial
 CC expression vector, and provides a high-efficiency recovery of
 CC recombinants for generating novel polynucleotides from parental templates

CC that are more diverse and with a lower percentage of sequence identity.
CC This sequence represents a tomato mosaic virus movement protein gene used
CC in the method of the invention

XX SQ Sequence 807 BP; 249 A; 121 C; 215 G; 222 T; 0 U; 0 Other;
Query Match 100.0%; Score 76; DB 8; Length 807;
Best Local Similarity 100.0%; Pred. No. 2.8e-16; Length 807;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGAGAGCGGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
Db 541 GTGAGAGCGGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 600
Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 6
ADM68440
ID ADM68440 standard; DNA; 807 BP.
XX AC ADM68440;
XX DT 03-JUN-2004 (first entry)
XX DE Tobacco mosaic virus movement protein gene #1.
XX KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
XX KW single nucleotide polymorphism; cancer susceptibility;
XX KW sequence variation redistribution; movement protein; gene.
XX OS Tobacco mosaic virus.
XX FN US2003157682-A1.
XX PD 21-AUG-2003.
XX PF 31-JAN-2003; 2003US-00356708.
XX PR 01-FEB-2002; 2002US-0353722P.
XX PR 14-MAR-2002; 2002US-00098155.
XX PR 01-AUG-2002; 2002US-00211079.
XX PA (PADG/) PADGETT H S.
XX PA (VAEW/) VAEWHONGS A A.
XX PA (VOJD/) VOJDANI F S.
XX PA (SMIT/) SMITH M L.
XX PA (LIND/) LINDBO J A.
XX PA (FITZ/) FITZMAURICE W P.
XX PI Padgett HS, Vaewhongs AA, Vojdani PS, Smith ML, Lindbo JA;
XX PI Fitzmaurice WP;
XX DR WPI; 2003-766176/72.

XX Making a mismatch endonuclease, useful in gene shuffling and in detection
XX of single nucleotide polymorphisms, comprises transfecting a host with a
XX recombinant viral vector including a polynucleotide encoding a mismatch
XX endonuclease.

XX Example 13; SEQ ID NO 9; 79pp; English.

XX The invention relates to a method of making a mismatch endonuclease
XX enzyme comprising transfecting a host plant, animal, yeast, fungus or
XX bacterium with a recombinant viral vector that encodes a polynucleotide
XX sequence for a mismatch endonuclease, growing the host so that the
XX polynucleotide is expressed, and extracting the mismatch endonuclease
XX enzyme from the host. The method is useful for making mismatch
XX endonuclease enzymes, for obtaining peptides and polynucleotides with
XX desired functional properties and for detecting mutations. The mismatch
XX endonuclease enzymes are useful in gene shuffling technology for

CC developing new genes, in detecting single nucleotide polymorphisms for
CC e.g. detecting evidence of cancer susceptibility, or in redistributing
CC sequence variations between non-identical polynucleotide sequences. The
CC present sequence represents a tobacco mosaic virus movement protein gene.
XX SQ Sequence 807 BP; 249 A; 121 C; 215 G; 222 T; 0 U; 0 Other;
Query Match 100.0%; Score 76; DB 11; Length 807;
Best Local Similarity 100.0%; Pred. No. 2.8e-16; Length 807;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGAGCGGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
Db 541 GTGAGAGCGGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 600
Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 7
ADP26593
ID ADP26593 standard; DNA; 807 BP.
XX AC ADP26593;
XX DT 26-AUG-2004 (first entry)
XX DE Tobamovirus DNA #1.
XX KW Sequence variation; heteroduplex; transcription; DNA integration;
XX KW ribozyme expression; gene; ds; tobamovirus.
XX OS Tobamovirus.
XX FN US2004110130-A1.
XX PD 10-JUN-2004.
XX PF 25-OCT-2002; 2002US-00280913.
XX PR 02-FEB-2001; 2001US-0266386P.
XX PR 14-FEB-2001; 2001US-0268785P.
XX PR 01-FEB-2002; 2002US-00066390.
XX PR 08-AUG-2002; 2002US-0402342P.
XX PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX PI Padgett HS, Lindbo JA, Fitzmaurice WP;
XX DR WPI; 2004-440326/41.

XX Redistributing sequence variations between non-identical polynucleotide
XX sequences, useful for generating improved polynucleotide having a desired
XX characteristic, comprises making a heteroduplex and introducing a nick.

XX Example 10; SEQ ID NO 9; 75pp; English.

XX The invention relates to an in vitro method of redistributing sequence
XX variations between non-identical polynucleotide sequences, comprising
XX making a heteroduplex polynucleotide from two non-identical
XX polynucleotides, introducing a nick in the second strand at or near a
XX base pair mismatch site, removing the mismatched base(s) from the
XX mismatch site where the nick occurred and using the first strand as a
XX template to replace the removed base(s) with bases that complement the
XX base(s) in the first strand. The invention also relates to an in vitro
XX method of making a population of sequence variants from a heteroduplex
XX polynucleotide sequence, obtaining a polynucleotide sequence encoding a
XX desired functional property and identifying a reassorted DNA molecule
XX encoding a protein with a desired functional property. The method is
XX useful for generating an improved polynucleotide sequence or a population
XX of improved polynucleotide sequences possessing at least one desired
XX phenotypic characteristic (e.g., promotes transcription of linked

CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.

XX
SQ Sequence 807 BP; 249 A; 121 C; 215 G; 222 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 12; Length 807;
Best Local Similarity 100.0%; Pred. No. 2.8e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATGGAAGAT 60
Db 541 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATGGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 8
ADQ88415
ID ADQ88415 standard; DNA; 807 BP.

XX
AC ADQ88415;

DT 18-NOV-2004 (first entry)

XX TMV-U1 movement protein (MP) gene.

DE Molecular biology; TMV; tobacco mosaic virus; movement protein; MP; ds.

XX
KW Tobacco mosaic virus.

OS
XX US2004142433-A1.

XX
XX 22-JUL-2004.

XX
XX 10-OCT-2003; 2003US-00684134.

XX
XX 02-FEB-2001; 2001US-0266386P.

XX
XX 14-FEB-2001; 2001US-0268785P.

XX
XX 01-FEB-2002; 2002US-00066390.

XX
XX 08-AUG-2002; 2002US-0402342P.

XX
XX 21-AUG-2002; 2002US-00226372.

XX
XX 25-OCT-2002; 2002US-00280913.

XX
XX 08-AUG-2003; 2003US-00637758.

XX
XX (PADG/) PADGETT H S.
XX
XX (FITZ/) FITZMAURICE W P.
XX
XX (LIND/) LINDBO J A.
XX
XX (VAEW/) VAETHONGS A A.
XX
XX (VOJD/) VOJDANI F S.
XX
XX (SMIT/) SMITH M L.

XX
XX Padgett HS, Fitzmaurice WP, Lindbo JA, Vaethongs AA, Vojdani FS;
XX
XX Smith ML;

XX
XX WPI; 2004-552565/53.

XX
XX Preparing variant polynucleotides having different nucleotide sequences
XX from at least two parent polynucleotides, useful in molecular biology, in
XX particular for generating populations of related nucleic acid molecules.

XX
XX Example 10; SEQ ID NO 9; 81pp; English.

XX
XX The invention relates to a method of preparing a variant polynucleotide
XX having a different nucleotide sequence from at least two parent
XX polynucleotides. The method involves preparing at least one heteroduplex
XX between the two parent polynucleotides, cleaving at least one
XX polynucleotide strand in the heteroduplex at a mismatch site to form a
XX cleavage site, replacing at least one nucleotide on at least one strand

CC at or near the cleavage site where at least one of the strands has a
CC different nucleotide sequence from either of the at least two parent
CC polynucleotides. Methods and compositions of the invention are useful in
CC the field of molecular biology in particular for generating populations
CC of related nucleic acid molecules. The present sequence is tobacco mosaic
CC virus U1 type strain (TMV-U1) movement protein (MP) gene. This sequence
CC is used in the exemplification of the invention

XX
SQ Sequence 807 BP; 249 A; 121 C; 215 G; 222 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 13; Length 807;

Best Local Similarity 100.0%; Pred. No. 2.8e-16;

Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATGGAAGAT 60
Db 541 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATGGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 9

ADSI7169

ID ADSI7169 standard; DNA; 807 BP.

XX
AC ADSI7169;

DT 02-DEC-2004 (first entry)

XX Tobacco mosaic virus U1 type strain movement protein (MP) gene.

XX
KW Polymerase; movement protein; MP; gene; ds.

XX
OS Tobacco mosaic virus.

XX
XX US2004180352-A1.

XX
XX 16-SEP-2004.

XX
XX 08-AUG-2003; 2003US-00637758.

XX
XX 08-AUG-2002; 2002US-0402342P.

XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.

XX
XX Padgett HS, Lindbo JA, Fitzmaurice WP, Vaethongs AA;

XX
XX WPI; 2004-667658/65.

XX
XX In vitro method of making sequence variants from heteroduplex
XX polynucleotide, involves combining polynucleotide with agent having
XX polymerase and strand cleavage activities for sufficient time for
XX percentage of complementarity to increase.

XX
XX Example 10; SEQ ID NO 9; 81pp; English.

XX
XX The invention relates to an in vitro method of making sequence variants
XX from heteroduplex polynucleotide which involves combining polynucleotide
XX with an agent having polymerase and strand cleavage activities for
XX sufficient time for the percentage of complementarity to be increased
XX within the heteroduplex. The method is useful for performing in vitro
XX method of making sequence variants from one or more heteroduplex
XX polynucleotide. It is useful in increasing diversity in a population of
XX sequences and obtaining a polynucleotide encoding a desired functional
XX property. The present sequence is the Tobacco mosaic virus U1 type strain
XX (TMV-U1) movement protein (MP) gene. This sequence is used to illustrate
XX the method of the invention.

XX
SQ Sequence 807 BP; 249 A; 121 C; 215 G; 222 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 13; Length 807;

RESULT 12
ADK56953
ID ADK56953 standard; DNA; 1187 BP.
XX
AC ADK56953;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plant DNA sequence which confers altered metabolic characteristic #4336.
XX
KW altered metabolic characteristic; plant; acid metabolism;
KW alcohol metabolism; fatty acid metabolism;
KW branched fatty acid metabolism; alkalioid metabolism;
KW amino acid metabolism; ester metabolism; glyceride metabolism;
KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
OS Unidentified.
XX
FN WO2003020936-A1.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027884.
XX
PR 31-AUG-2001; 2001US-0316471P.
XX
PA (DOWC) DOW CHEM CO.
PA (DOWC) DOW AGROSCIENCES LLC.
XX
XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
PI Orledo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX WPI; 2003-313091/30.
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
PT benthamiana plants, useful for altering the levels of metabolites e.g.
PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
PS Claim 1; SEQ ID NO 4336; 2576pp; English.
XX
CC The invention comprises DNA sequences which confer an altered metabolic
CC characteristic when they are expressed in a plant. The DNA sequences of
CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkalioid or other
CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC altered glyceride metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.
XX
SQ Sequence 1187 BP; 379 A; 210 C; 274 G; 324 T; 0 U; 0 Other;
Query Match 100.0%; Score 76; DB 10; Length 1187;
Best Local Similarity 100.0%; Pred. No. 3.1e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGAGAGACGGAGGCCCATGGAACCTTACAGAAGAGTCGTTGATGAGTTCATGAAGAT 60
Db 309 GTGAGAGACGGAGGCCCATGGAACCTTACAGAAGAGTCGTTGATGAGTTCATGAAGAT 368
Qy 61 GTCCCTATGTCGATCA 76
Db 369 GTCCCTATGTCGATCA 384

RESULT 13
AAQ67663
ID AAQ67663 standard; RNA; 1825 BP.
XX
AC AAQ67663;
XX
DT 25-MAR-2003 (revised)
DT 15-FEB-1995 (first entry)
XX
DE TMV replicon RNA.
XX
KW Replicon; transgenic plant; tobacco; Nicotiana tabacum; helper virus;
KW RNA virus; chloramphenicol-acetyltransferase; CAT; movement protein; TMV;
KW replication; ss.
OS Tobacco mosaic virus.
XX
FH Key Location/Qualifiers
FT CDS 70..876
FT /*tag= a
FT /product= "TMV 30 kDa movement protein"
FT 916..1530
FT /*tag= b
FT /product= "CAT"
XX
FN WO9416089-A1.
XX
PD 21-JUL-1994.
XX
PF 29-DEC-1993; 93WO-US012636.
XX
PR 30-DEC-1992; 92US-00997733.
XX
PA (BIOS-) BIOSOURCE GENETICS CORP.
XX
PI Turpen TH;
XX
DR WPI; 1994-249235/30.
DR P-PSDB; AAR56447.
XX
XX Viral amplification of recombinant mRNA in transgenic plants esp. tobacco
PT - by using replicons having replication function regulated by helper
PT virus.
XX
PS Disclosure; Page 38-44; 60pp; English.
XX
CC A sequence is provided for the high level expression of chloramphenicol-
CC acetyltransferase (CAT) in tobacco plants by replicon RNA amplification
CC with helper virus and a movement protein gene from tobacco mosaic virus.
CC The sequence of the replicon RNA, produced by host transcription, RNA
CC processing and replication in the presence of helper virus, is given in
CC AAQ67663. (Updated on 25-MAR-2003 to correct FN field.)
XX
SQ Sequence 1825 BP; 543 A; 341 C; 437 G; 0 T; 504 U; 0 Other;
Query Match 100.0%; Score 76; DB 2; Length 1825;
Best Local Similarity 77.6%; Pred. No. 3.5e-16;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGAGAGACGGAGGCCCATGGAACCTTACAGAAGAGTCGTTGATGAGTTCATGAAGAT 60
Db 610 GUGAGAGACGGAGGCCCATGGAACCTTACAGAAGAGTCGTTGATGAGTTCATGAAGAT 669
Qy 61 GTCCCTATGTCGATCA 76
Db 670 GUCCCUAUGUGCAUCA 685
RESULT 14
AAV54825
ID AAV54825 standard; RNA; 1825 BP.
XX
AC AAV54825;

Search completed: September 26, 2005, 01:44:52
Job time : 97.5556 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 00:52:04 ; Search time 615.444 Seconds
(without alignments)
4700.480 Million cell updates/sec

Title: US-09-551-494-2_COPY_5443_5518

Perfect score: 76
Sequence: 1 gtgagagacgagggcccat.....agatgtccctatgtcgatca 76

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gsl1:*
- 9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	45.6	60.0	188	4	BM068137 KS08017C0
C 2	45.6	60.0	348	4	BM067518 KS08006E1
C 3	31.8	41.8	827	9	BM180131 Danilo rer
C 4	30.6	40.3	626	3	BM288685 Arabidops
C 5	30	39.5	721	8	CC430046 PUH7H76TB
C 6	30	39.5	929	9	CG045306 PUJWB66TD
C 7	30	39.5	933	9	CG244112 OGWBI28TV
C 8	29.8	39.2	606	4	BJ685374 BJ685374
C 9	29.6	38.9	1039	8	BZ466235 BOKRM77TR
C 10	29.2	38.4	480	8	BH279563 CH230-128
C 11	29.2	38.4	482	1	AA728539 32948 Lam
C 12	29.2	38.4	494	7	CO174153 NDLI 42 C
C 13	29.2	38.4	519	7	CO365231 RTK1 24 A
C 14	29.2	38.4	520	8	BH284760 BH230-128
C 15	29.2	38.4	620	4	BU701812 BU701812
C 16	28.6	37.6	615	9	BM180239 Danilo rer
C 17	28.6	37.6	784	5	BK304531 BK304531
C 18	28.4	37.4	431	2	AW616937 EST323348
C 19	28.4	37.4	501	5	BQ858541 QGC10112
C 20	28.4	37.4	667	5	BU013451 QGJAL01 Y
C 21	28.4	37.4	685	8	BZ977915 PUGX74TB
C 22	28.2	37.1	239	9	BX942949 Arabidops
C 23	28.2	37.1	392	1	AV533338 AV533338
C 24	28.2	37.1	464	7	T04811 859 Lambda-

25	28.2	37.1	645	6	CA766726	CA766726	AF53-Rpf-
C 26	28.2	37.1	661	1	AV822458	AV822458	AV822458
C 27	28.2	37.1	910	5	BX426792	BX426792	BX426792
C 28	28.2	37.1	1240	3	CNS0A3MT	CNS0A3MT	Arabidops
C 29	28	36.8	125	5	BQ875013	BQ875013	OGI6015 Y
C 30	28	36.8	432	1	AV546240	AV546240	AV546240
C 31	28	36.8	497	6	CB260582	CB260582	84-E9409-
C 32	28	36.8	603	7	W78405	W78405	me78h06.r1
C 33	28	36.8	688	4	BI955206	BI955206	HVSMEM002
C 34	27.8	36.6	438	8	AQ685535	AQ685535	HS 5528 B
C 35	27.8	36.6	456	8	BH025136	BH025136	RPCL-24-2
C 36	27.8	36.6	587	9	DR21K14T	DR21K14T	Danilo rer
C 37	27.8	36.6	947	1	AL543898	AL543898	AL543898
C 38	27.8	36.6	1047	8	CC174831	CC174831	ZMMBB029
C 39	27.8	36.6	2327	3	CR625340	CR625340	full-length
C 40	27.6	36.3	203	8	BH191265	BH191265	TC3-40C14
C 41	27.6	36.3	302	8	BH190855	BH190855	TC3-37119
C 42	27.6	36.3	330	8	BH191345	BH191345	TC3-38823
C 43	27.6	36.3	371	8	BH190864	BH190864	TC3-38C24
C 44	27.6	36.3	414	1	AU301905	AU301905	AU301905
C 45	27.6	36.3	469	8	BH842029	BH842029	TC3-53D6.

ALIGNMENTS

RESULT 1
BM068137/c 188 bp mRNA linear EST 11-SEP-2002
LOCUS BM068137 KS08017C07 KS08 Capsicum annuum cDNA, mRNA sequence.
DEFINITION BM068137
ACCESSION BM068137.1 GI:22788242
VERSION EST.
KEYWORDS
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Capsicum.
REFERENCE 1 (bases 1 to 188)
AUTHORS Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S., Hur,C.-G. and Choi,D.
TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen
JOURNAL Unpublished (2001)
COMMENT Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doilemail.kribb.re.kr
High quality sequence stop: 188.
Location/Qualifiers
source 1..188
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stage="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"

ORIGIN

Query Match 60.0%; Score 45.6; DB 4; Length 188;
Best Local Similarity 75.0%; Pred. No. 0.00017;
Matches 57; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 GTGAGACGAGGCGCCCATCGAAGTTCACAGAGTCGTTGATGAGTTCATCGAAGT 60
DB 122 GTGTGCGAAGGAGGACCGGTGAACCTTACAGAGCAGTTGTTGATGAGTTCATCGAATCA 63

Qy 61 GTCCCTATGTCGATCA 76
 Db 62 GTTCCAAATGGCTGACA 47

RESULT 2
 BM067518/c
 LOCUS BM067518 348 bp mRNA linear EST 11-SEP-2002
 DEFINITION KS08006E10 KS08 Capsicum annuum cDNA, mRNA sequence.
 ACCESSION BM067518
 VERSION BM067518.1 GI:22787638
 KEYWORDS EST.
 SOURCE Capsicum annuum
 ORGANISM Capsicum annuum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Capsicum.
 REFERENCE 1 (bases 1 to 348)
 AUTHORS Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S., Hur, C.-G. and Choi, D.
 TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen
 JOURNAL Unpublished (2001)
 COMMENT Contact: Doil Choi
 Genome Research Center and National Center for Genome Information
 Korea Research Institute of Bioscience and Biotechnology
 P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
 Tel: 82-42-860-4340
 Fax: 82-42-860-4309
 Email: doilemail.kribb.re.kr
 High quality sequence stop: 348.
 Location/Qualifiers
 1..348
 /organism="Capsicum annuum"
 /mol_type="mRNA"
 /cultivar="Hang Keun"
 /db_xref="taxon:4072"
 /tissue_type="anther"
 /dev_stage="10 weeks after germination"
 /clone_lib="KS08"
 /note="Vector: pluescript SK(-)"

Query Match 60.0%; Score 45.6; DB 4; Length 348;
 Best Local Similarity 75.0%; Pred. No. 0.00019;
 Matches 57; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GTGAGACGGAGGCCCATGGAACCTTACAGAGAAGTCGTTGATGATGATGGAAGAT 60
 Db 294 GTGTCGGAAGGAGGACCGGTTGAACCTTACAGAGAAGTCGTTGATGATGATGCAATCA 235

Qy 61 GTCCCTATGTCGATCA 76
 Db 234 GTTCCAAATGGCTGACA 219

RESULT 3
 BX180131
 LOCUS BX180131 827 bp DNA linear GSS 28-JAN-2003
 DEFINITION Danio rerio genomic clone DKEY-172D22, genomic survey sequence.
 ACCESSION BX180131
 VERSION BX180131.1 GI:28011935
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 827)
 AUTHORS Humphray, S.J., Huckle, R. and Durham, J.L.
 TITLE Direct Submission
 JOURNAL Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome

Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished from the T7 end of BAC 172D22. 172D22 is part of the Daniokey BAC Library created by R. Piasterk and N.V. Keygene. Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

Location/Qualifiers
 1..827
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEY-172D22"
 /tissue_type="Testis"
 /note="vector pindigoBAC-536"

Query Match 41.8%; Score 31.8; DB 9; Length 827;
 Best Local Similarity 76.5%; Pred. No. 8.7;
 Matches 39; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 14 GGCCCATGGAACCTTACAGAGAAGTCGTTGATGATGATGGAAGATGTC 64
 Db 88 GGCCCATGGAATTCGAATGAAGAAGTCGCTGATGAGATCTGGGAAGATCC 138

RESULT 4
 CNS0A316/c
 LOCUS CNS0A316 626 bp mRNA linear HTC 06-FEB-2004
 DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLT5IL22B02 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).
 ACCESSION BX828685
 VERSION BX828685.1 GI:42461132
 KEYWORDS HTC; GSLT cDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 626)
 AUTHORS Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 626)
 GENOSCOPE Genoscope.
 DIRECT SUBMISSION Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIFS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers
 1..626
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLT5IL22B02"

FEATURES
 source
 Location/Qualifiers
 1..626
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLT5IL22B02"

VERSION	CG045306.1	GI:33917486
KEYWORDS	GSS.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
AUTHORS	1 (bases 1 to 929) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Wiguell,P., Ma,J. and Bennettzen,J.	
TITLE	Maize Genomics Consortium	
JOURNAL	Unpublished (2003)	
COMMENT	Other_GSSs: PUBW66TB Contact: Cathy Whitelaw TIGR	
FEATURES	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends.	
source	Location/Qualifiers	
	1..929	
	/organism="Zea mays"	
	/mol_type="genomic DNA"	
	/strain="B73"	
	/db_xref="taxon:4577"	
	/clone="ZMMBT0633K11"	
	/clone_lib="ZM 0.6 1.0 KB"	
	/note="Vector: PCR4-ROPO; Site_1: EcoRI; 0.6-1.0 kb high CoT selected genomic DNA library"	
ORIGIN		
Query Match	39.5%;	Score 30; DB 9; Length 929;
Best Local Similarity	78.3%;	Pred: No. 35;
Matches	36; Conservative	0; Mismatches 10; Indels 0; Gaps 0
Qy	22 GAACCTACAGAGAAAGTCGTTGATGAGTTCATGGAAGATGTCCTA	67
Db		
	30 GACCTTATGAGAGTGGTTGATGTGATGATGGAAGGTGTCATTA	75
RESULT 7		
CG244112/c	933 bp	DNA linear
LOCUS	CGWB12FTV ZM_0.7_1.5_KB	Zea mays genomic clone ZMMBMA0327F07,
DEFINITION	genomic survey sequence.	
ACCESSION	CG244112	
VERSION	CG244112.1	GI:34143998
KEYWORDS	GSS.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
AUTHORS	1 (bases 1 to 933) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.	
TITLE	Consortium for Maize Genomics	
JOURNAL	Unpublished (2002)	
COMMENT	Other_GSSs: CGWBI28TH Contact: Cathy Whitelaw TIGR	
FEATURES	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends.	
source	Location/Qualifiers	
	1..933	

Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/or ering_information.htm). BAC end
 page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
 Plate: 128 row: D column: 22
 Seq primer: T7
 Class: BAC ends.

FEATURES

source
 Location/Qualifiers
 1. .480

/organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SeNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-128D22"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 1"
 /note="Vector: pTARBAC2.1; Site 1: EcorI; Site 2: EcorI;
 CHORI-230 Rat (BN/SeNHsd/MCW) BAC library produced by
 Pieter de Jong"

ORIGIN

Query Match 38.4%; Score 29.2; DB 8; Length 480;
 Best Local Similarity 62.2%; Pred. No. 57;
 Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 3 GAGAGCGAGCGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGATGT 62
 |||||
 Db 264 GAGAAAGCGAGGTCAAAGGCAGATGCATCATCAAGTCATTACTCATGTTAAGTAAATGT 323
 |||||
 Qy 63 CCCTATGTCGATCA 76
 |||||
 Db 324 CTCTCTGGCAACA 337
 |||||

RESULT 11
 AA728539/c
 LOCUS
 DEFINITION 32948 Lambda-PRL2 Arabidopsis thaliana cdna clone 32D8TT, mRNA
 sequence.
 AA728539
 AA728539.1 GI:2747496
 EST.
 Arabidopsis thaliana (chale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 482)
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
 McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
 Retzel,E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cdna clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 7846151

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@ibm.cl.msu.edu
 Seq primer: T7.

FEATURES

source
 Location/Qualifiers
 1. .482

/organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="32D8TT"
 /clone_lib="Lambda-PRL2"

ORIGIN

Query Match 38.4%; Score 29.2; DB 1; Length 482;
 Best Local Similarity 63.2%; Pred. No. 57;
 Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 8 ACGGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGATGTCCCTA 67
 |||||
 Db 365 ATGGAGAAGACATCGAATCTCCCGATGAAGTGAATGATGATGATGGAGATGCCAATG 306
 |||||
 Qy 68 TGTCGATC 75
 |||||
 Db 305 TGGAGATC 298
 |||||

RESULT 12
 COL174153
 LOCUS
 DEFINITION NDLL1_42_C08_b1_A029 Needles control Pinus taeda cdna clone
 COL174153
 COL174153.1 GI:48947025
 EST.
 Pinus taeda (loblolly pine)
 Pinus taeda
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

An EST database from untreated loblolly pine (Pinus taeda) needles
 Unpublished (2004)
 Other ESTs: NDLL1_42_C08.g1_A029
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz (School of
 Forest Resources, University of Georgia); plant material prepared
 by Craig Zimmermann (School of Forest Resources, University of
 Georgia) using rooted cuttings provided by the Forest Biology
 Research Cooperative (FBRC) and the CCLONES project a the
 University of Florida; sequencing done in the Laboratory for
 Genomics and Bioinformatics, University of Georgia. Sequence ends
 have been trimmed to exclude vector and regions below Phred quality
 16. Three-prime sequences are presented as their reverse complement
 and have been trimmed to exclude polyA.
 Seq primer: M13-21 (TGTAACAACGACGCCAGT)
 POLYA=Yes.

FEATURES

source
 Location/Qualifiers
 1. .494

/organism="Pinus taeda"
 /mol_type="mRNA"
 /strain="3 CCLONES"
 /db_xref="taxon:3352"
 /clone="NDLL1_42_C08_A029"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /clone_lib="Needles control"
 /note="Organ: needle; Vector: pSL1180; Site 1: EcoRI;
 Site 2: XhoI; The library was prepared from polyA+ RNA

/note="Vector: lambda zip-Lox; Site 1: Sal; Site 2: Not;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light; 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques). The vector is BRL's lambda Zip-Lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dt primed cDNA. "

from the needles of 1-year-old loblolly pine (*Pinus taeda*) cuttings that were rooted and then planted in washed sand. Just before harvesting needles for RNA isolation, the rooted cuttings were maintained for 27 days (April 2003) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

```

ORIGIN
Query Match      38.4%; Score 29.2; DB 7; Length 494;
Best Local Similarity 65.2%; Pred. No. 57;
Matches 43; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 8 ACGGAGGCCCCATCGAATTACAGAGAAGTCGTTGATGATTCATCGAAGATGCCCTA 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 217 ATGACGGCCCATGGCAAAAATAGCAGAAATAGAAATAGGTGTTTATAGCAAGGTGTTCTG 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 68 TGTGCA 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 277 TGTCTA 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
CO365231/c
LOCUS      CO365231      519 bp      mRNA      linear      EST 29-JUN-2004
DEFINITION RTK1_24_A11_g1_A029 Roots minus potassium Pinus taeda cDNA clone
            RTK1_24_A11_A029 5', mRNA sequence.
ACCESSION  CO365231
VERSION    CO365231.1 GI:49446548
KEYWORDS   EST.
SOURCE     Pinus taeda (loblolly pine)
ORGANISM   Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE  1 (bases 1 to 519)
AUTHORS   Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Zimmermann,C. and
            Dean,J.F.D.
TITLE     An EST database from potassium-deficient loblolly pine (Pinus
            taeda) roots
JOURNAL   Unpublished (2004)
COMMENT   Other_ESTs: RTK1_24_A11.b1_A029
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            RNA prepared and library constructed by W. Walter Lorenz (School of
            Forest Resources, University of Georgia); plant material prepared
            by Craig Zimmermann (School of Forest Resources, University of
            Georgia) using rooted cuttings provided by the Forest Biology
            Research Cooperative (FBRC) and the CCLONES project a the
            University of Florida; sequencing done in the laboratory for
            Genomics and Bioinformatics, University of Georgia. Sequence ends
            have been trimmed to exclude vector and regions below Phred quality
            16. Three-prime sequences are presented as their reverse complement
            and have been trimmed to exclude polyA.
            Seq primer: JENREV (CAGGAACAGCTATGACC).
            Location/Qualifiers
                1. .519
                   /organism="Pinus taeda"
                   /mol_type="mRNA"
                   /strain="3 CCLONES"
                   /db_xref="taxon:3352"
                   /clone="RTK1_24_A11_A029"
                   /lab_host="DH10B-T1 phage-resistant E. coli"
                   /clone_lib="Roots minus potassium"
                   /note="Organ: Root; Vector: pSL1180; Site_1: EcoRI;

```

Site 2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (*Pinus taeda*) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 117 days (July 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. For twenty-eight days (28 d) prior to harvesting roots for mRNA preparation, the trees received Hoagland's solution lacking potassium (K) to induce a potassium-deficiency. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

```

ORIGIN
Query Match      38.4%; Score 29.2; DB 7; Length 519;
Best Local Similarity 65.2%; Pred. No. 58;
Matches 43; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 8 ACGGAGGCCCCATCGAATTACAGAGAAGTCGTTGATGATTCATCGAAGATGCCCTA 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 201 ATGACGGCCCATGGCAAAAATAGCAGAAATAGTATAGGTGTTTATAGCAAGGTGTTCTG 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 68 TGTGCA 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 TGTCTA 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
BH284760
LOCUS      BH284760      520 bp      DNA      linear      GSS 30-NOV-2001
DEFINITION CH230-128P14-TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-128P14, genomic survey sequence.
ACCESSION  BH284760.1 GI:17197166
VERSION    BH284760
KEYWORDS   GSS.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 520)
AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shivatsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.
            Rat BAC End Sequences from Library CHORI-230 EcoRI segment
            Unpublished (1999)
            Other GSSs: CH230-128P14.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.choi.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/oreringinformation.htm). BAC end
            page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html
            Plate: 128 Row: P Column: 14
            Seq primer: T7
            Class: BAC ends.
            Location/Qualifiers
                1. .520
                   /organism="Rattus norvegicus"
                   /mol_type="genomic DNA"
                   /strain="BN/SENHsd/MCW"
                   /db_xref="taxon:10116"
                   /clone="CH230-128P14"
                   /sex="Female"

```

FEATURES
source

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 01:18:09 ; Search time 29.6667 Seconds
(without alignments)
4191.810 Million cell updates/sec

Title: US-09-551-494-2_COPY_5443_5518

Perfect score: 76
Sequence: 1 gtgagacgagggcccat.....agatgtccctatgtcgatca 76

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	433	1	US-07-971-101-5
2	76	100.0	807	4	US-09-565-616A-3
3	76	100.0	807	4	US-09-565-616A-4
4	76	100.0	1825	1	US-08-176-414B-1
5	76	100.0	1825	2	US-08-336-724-1
6	76	100.0	6395	2	US-08-687-559-2
7	76	100.0	6395	3	US-09-259-741-1
8	76	100.0	6395	3	US-09-037-751-1
9	76	100.0	6395	3	US-09-466-422-1
10	76	100.0	6395	4	US-09-401-415-2
11	76	100.0	6395	4	US-09-962-527-1
12	76	100.0	6425	3	US-09-259-741-3
13	76	100.0	6425	3	US-09-037-751-3
14	76	100.0	6425	3	US-09-466-422-3
15	76	100.0	6425	4	US-09-962-527-3
16	76	100.0	6439	3	US-09-259-741-2
17	76	100.0	6439	3	US-09-037-751-2
18	76	100.0	6439	3	US-09-466-422-2
19	76	100.0	6439	4	US-09-962-527-2
20	76	100.0	6446	3	US-09-259-741-5
21	76	100.0	6446	3	US-09-037-751-5
22	76	100.0	6446	3	US-09-466-422-5
23	76	100.0	6446	4	US-09-962-527-5
24	76	100.0	6475	3	US-09-259-741-4
25	76	100.0	6475	3	US-09-037-751-4
26	76	100.0	6475	3	US-09-466-422-4
27	76	100.0	6475	4	US-09-962-527-4

28	76	100.0	7685	3	US-09-502-710-22	Sequence 22, Appl
29	76	100.0	7685	3	US-09-502-710-25	Sequence 25, Appl
30	76	100.0	7685	3	US-09-502-711-22	Sequence 22, Appl
31	76	100.0	7685	3	US-09-502-711-25	Sequence 25, Appl
32	76	100.0	7685	4	US-09-565-616A-1	Sequence 1, Appl
33	76	100.0	7686	3	US-09-502-710-23	Sequence 23, Appl
34	76	100.0	7686	3	US-09-502-710-26	Sequence 26, Appl
35	76	100.0	7686	3	US-09-502-711-23	Sequence 23, Appl
36	76	100.0	7686	4	US-09-502-711-26	Sequence 26, Appl
37	76	100.0	7686	4	US-09-565-616A-2	Sequence 2, Appl
38	76	100.0	7687	3	US-09-502-710-24	Sequence 24, Appl
39	76	100.0	7687	3	US-09-502-711-24	Sequence 27, Appl
40	76	100.0	7688	3	US-09-502-710-27	Sequence 27, Appl
41	76	100.0	7688	3	US-09-502-711-27	Sequence 1, Appl
42	76	100.0	7926	3	US-09-500-554-1	Sequence 1, Appl
43	76	100.0	7926	3	US-09-726-648-1	Sequence 1, Appl
44	76	100.0	7926	4	US-10-119-330-1	Sequence 1, Appl
45	75	98.7	75	1	US-07-971-101-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-07-971-101-5
; Sequence 5, Application US/07971101
; Patent No. 5443969
; GENERAL INFORMATION:
; APPLICANT: Wilson, Thomas M.A. et al.
; TITLE OF INVENTION: RNA Packaging System
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/971,101
; FILING DATE: 19921029
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7108-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 nucleotides
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: RNA
US-07-971-101-5

Query Match 100.0%; Score 76; DB 1; Length 433;
Best Local Similarity 77.6%; Pred. No. 3.1e-17;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGAGACGAGGGCCCATGGAAGTTCAGAACGTCGTTGATGAGTTCATCGAGAT 60
Db 326 GUGAGACGAGGGGGCCCAUGGAACUACAGAGGUCGUUGAUGAGUUCAGGAAGAU 385
Qy 61 GTCCTATGTCGATCA 76

```
Db      386 GUCCUAUGUGCAUCA 401
      |:|:|:|:|:|:|:|:|:|
RESULT 2
US-09-565-616A-3
; Sequence 3, Application US/09565616A
; Patent No. 6656726
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P.
; APPLICANT: Pogue, Gregory P.
; APPLICANT: Lindbo, John A.
; TITLE OF INVENTION: VIRAL EXPRESSION VECTORS
; FILE REFERENCE: 008010165US01
; CURRENT APPLICATION NUMBER: US/09/565,616A
; CURRENT FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60,132,697
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-09-565-616A-3
Query Match      100.0%; Score 76; DB 4; Length 807;
Best Local Similarity 100.0%; Pred. No. 3.8e-17;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 GTGAGACGCGAGGGCCCATCGAACTTACAGAAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db      541 GTGAGACGCGAGGGCCCATCGAACTTACAGAAAGTCGTTGATGAGTTTCATGGAAGAT 600
Qy      61 GTCCCTATGTCGATCA 76
Db      601 GTCCCTATGTCGATCA 616
RESULT 3
US-09-565-616A-4
; Sequence 4, Application US/09565616A
; Patent No. 6656726
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P.
; APPLICANT: Pogue, Gregory P.
; APPLICANT: Lindbo, John A.
; TITLE OF INVENTION: VIRAL EXPRESSION VECTORS
; FILE REFERENCE: 008010165US01
; CURRENT APPLICATION NUMBER: US/09/565,616A
; CURRENT FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60,132,697
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-09-565-616A-4
Query Match      100.0%; Score 76; DB 4; Length 807;
Best Local Similarity 100.0%; Pred. No. 3.8e-17;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 GTGAGACGCGAGGGCCCATCGAACTTACAGAAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db      541 GTGAGACGCGAGGGCCCATCGAACTTACAGAAAGTCGTTGATGAGTTTCATGGAAGAT 600
Qy      61 GTCCCTATGTCGATCA 76
Db      601 GTCCCTATGTCGATCA 616
```

```
RESULT 4
US-08-176-414B-1
; Sequence 1, Application US/08176414B
; Patent No. 5811653
; GENERAL INFORMATION:
; APPLICANT: Turpen, Thomas H.
; TITLE OF INVENTION: VIRAL AMPLIFICATION OF RECOMBINANT
; TITLE OF INVENTION: MESSENGER RNA IN TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington D.C.
; STATE: District of Columbia
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,414B
; FILING DATE: 29-Dec-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 00801.0038.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: RNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 70..1527
; US-08-176-414B-1
Query Match      100.0%; Score 76; DB 1; Length 1825;
Best Local Similarity 77.6%; Pred. No. 5e-17;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;
Qy      1 GTGAGACGCGAGGGCCCATCGAACTTACAGAAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db      610 GUGAGACGCGAGGGCCCAUGGAACTTACAGAAAGTCGTTGATGAGTTTCATGGAAGAT 600
Qy      61 GTCCCTATGTCGATCA 76
Db      670 GUCCUAUGUGCAUCA 685
RESULT 5
US-08-336-724-1
; Sequence 1, Application US/08336724
; Patent No. 5965794
; GENERAL INFORMATION:
; APPLICANT: Turpen, Thomas H.
; TITLE OF INVENTION: VIRAL AMPLIFICATION OF
; TITLE OF INVENTION: RECOMBINANT MESSENGER RNA IN TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CAL
```

```
;
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,724
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/997,733
; FILING DATE: 30-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: BIOG-20220 USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-433-4150
; TELEFAX: 415-433-8716
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1825
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (episomal), peptide
; DESCRIPTION: Peptide encodes for TMV 30kDa
; DESCRIPTION: movement protein (268 residues) and CAT (204
; DESCRIPTION: residues).
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tobacco Mosaic Virus
; IMMEDIATE SOURCE:
; CLONE:
; FEATURE:
;
; US-08-336-724-1
```

```
Query Match 100.0%; Score 76; DB 2; Length 1825;
Best Local Similarity 77.6%; Pred. No. 5e-17;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db 610 GUGAGACGCGAGGCGCCCAUGGAACUUAACAGAGAAGUGUGUUGAUGAUGAUGAUGAUGA 669

Qy 61 GTCCCTATGTCGATCA 76
Db 670 GUCCCUAUGUGCAUCA 685
```

```
RESULT 6
US-08-687-559-2
; Sequence 2, Application US/08687559
; Patent No. 5955647
; GENERAL INFORMATION:
; APPLICANT: Fitchen, John H.
; APPLICANT: Beachy, Roger N.
; TITLE OF INVENTION: METHOD FOR USING TOBACCO MOSAIC VIRUS TO
; TITLE OF INVENTION: OVERPRODUCE PEPTIDES AND PROTEINS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
```

```
;
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,559
; FILING DATE: No. 5955647ember 18, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01467
; FILING DATE: 03-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Learn, June M.
; REGISTRATION NUMBER: 31,238
; REFERENCE/DOCKET NUMBER: 07302/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: TMV
;
; US-08-687-559-2
```

```
Query Match 100.0%; Score 76; DB 2; Length 6395;
Best Local Similarity 100.0%; Pred. No. 7.6e-17;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db 5443 GTGAGACGCGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 5502

Qy 61 GTCCCTATGTCGATCA 76
Db 5503 GTCCCTATGTCGATCA 5518
```

```
RESULT 7
US-09-259-741-1
; Sequence 1, Application US/09259741
; Patent No. 6033895
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: MCCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
; TITLE OF INVENTION: SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/259,741
; FILING DATE: February 25, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: March 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
```



```

Db      5443  GUGAGACGCGGGCCCAUGGAAGACTUACAGAGAGUCGUGAUGAGUUAUGGAAGAU 5503
Qy      61  GTCCTATGTCGATCA 76
          |:|:|:|:|:|:|:|:|
Db      5503  GUCCUAUGUCGAUCA 5518

RESULT 10
US-09-401-415-2
; Sequence 2, Application US/09401415
; Patent No. 6503732
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; TITLE OF INVENTION: METHOD FOR USING TOBACCO MOSAIC VIRUS TO
; OVERPRODUCE PEPTIDES AND PROTEINS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/401,415
; FILING DATE: 21-Sep-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01467
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bostich, June M.
; REGISTRATION NUMBER: 31,238
; REFERENCE/DOCKET NUMBER: FD-4074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: TWV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..6395
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-401-415-2

Query Match      100.0%; Score 76; DB 4; Length 6395;
Best Local Similarity 100.0%; Pred. No. 7.6e-17;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  GTGAGACGCGAGGGCCCATGGAACTTACAGAAGAGTCGTTGATGAGTTTCATGGAAGAT 60
Db      5443  GTGAGACGCGAGGGCCCATGGAACTTACAGAAGAGTCGTTGATGAGTTTCATGGAAGAT 60
Qy      61  GTCCTATGTCGATCA 76
          |:|:|:|:|:|:|:|:|
Db      5503  GUCCUAUGUCGAUCA 5518

RESULT 11
US-09-962-527-1
; Sequence 1, Application US/09962527

```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/259,741
; FILING DATE: February 25, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: March 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8100
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6425 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; US-09-259-741-3

Query Match 100.0%; Score 76; DB 3; Length 6425;
Best Local Similarity 77.6%; Pred.No. 7.6e-17;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATGGAACCTTACAGAAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5443 GUGAGACGCGAGGGCCCAUGGACUACAGAGAAGUCGUUGAUGAGUUCAGGAGAGAU 5502
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 GTCCTATGTCGATCA 76
|:|||||:|||||:
Db 5503 GUCCCUAUGUGCAUCA 5518

RESULT 13
US-09-037-751-3
; Sequence 3, Application US/09037751
; Patent No. 6037456
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: MCCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

```
;
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,751
; FILING DATE: 10-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6425 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; US-09-037-751-3

Query Match 100.0%; Score 76; DB 3; Length 6425;
Best Local Similarity 77.6%; Pred.No. 7.6e-17;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATGGAACCTTACAGAAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5443 GUGAGACGCGAGGGCCCAUGGACUACAGAGAAGUCGUUGAUGAGUUCAGGAGAGAU 5502
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 GTCCTATGTCGATCA 76
|:|||||:|||||:
Db 5503 GUCCCUAUGUGCAUCA 5518

RESULT 14
US-09-466-422-3
; Sequence 3, Application US/09466422
; Patent No. 6303779
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: MCCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; FROM PLANT SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/466,422
; FILING DATE: 17-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
```

```
;
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6425 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-466-422-3

Query Match      100.0%; Score 76; DB 3; Length 6425;
Best Local Similarity 77.6%; Pred. No. 7.6e-17;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
Db 5443 GUGAGACGCGAGGGCCCAUGGAACUACAGAGAAGCUGUGAUGAGUUCAGUGAAGAU 5502

Qy 61 GTCCTATGTCGATCA 76
Db 5503 GUCCCUAUGUGCAUCA 5518

Search completed: September 26, 2005, 04:42:15
Job time : 38.6667 secs
```

```
;
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6425 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-466-422-3

Query Match      100.0%; Score 76; DB 3; Length 6425;
Best Local Similarity 77.6%; Pred. No. 7.6e-17;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
Db 5443 GUGAGACGCGAGGGCCCAUGGAACUACAGAGAAGCUGUGAUGAGUUCAGUGAAGAU 5502

Qy 61 GTCCTATGTCGATCA 76
Db 5503 GUCCCUAUGUGCAUCA 5518

RESULT 15
US-09-962-527-3
; Sequence 3, Application US/09962527
; Patent No. 6740740
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; HOLTZ, R. BARRY
; MCCULLOCH, MICHAEL
; TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; FROM PLANT SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,527
; FILING DATE: 24-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: 10-march-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6425 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
;
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-962-527-3

Query Match      100.0%; Score 76; DB 4; Length 6425;
Best Local Similarity 77.6%; Pred. No. 7.6e-17;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTTCATCGAAGAT 60
Db 5443 GUGAGACGCGAGGGCCCAUGGAACUACAGAGAAGCUGUGAUGAGUUCAGUGAAGAU 5502

Qy 61 GTCCTATGTCGATCA 76
Db 5503 GUCCCUAUGUGCAUCA 5518

Search completed: September 26, 2005, 04:42:15
Job time : 38.6667 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 01:30:34 ; Search time 134 Seconds
(without alignments)
3792.150 Million cell updates/sec

Title: US-09-551-494-2_COPY_5443_5518

Perfect score: 76

Sequence: 1 gtgagagacggagggcccat.....agatgtccctatgtcgatca 76

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
c 1	76	100.0	557	21	US-10-487-901-4335 Sequence 4335, Ap
2	76	100.0	807	10	US-09-775-049-17 Sequence 17, Appl
3	76	100.0	807	13	US-10-066-390-10 Sequence 10, Appl
4	76	100.0	807	13	US-10-206-030-10 Sequence 10, Appl
5	76	100.0	807	15	US-10-211-079-10 Sequence 10, Appl
6	76	100.0	807	16	US-10-356-708-9 Sequence 9, Appl
7	76	100.0	807	16	US-10-205-772-10 Sequence 10, Appl

8	76	100.0	807	19	US-10-280-913A-9 Sequence 9, Appl
9	76	100.0	807	19	US-10-684-134-9 Sequence 9, Appl
10	76	100.0	807	19	US-10-637-758-9 Sequence 9, Appl
11	76	100.0	807	22	US-10-226-372A-10 Sequence 10, Appl
12	76	100.0	807	22	US-10-624-193-3 Sequence 3, Appl
13	76	100.0	807	22	US-10-624-193-4 Sequence 4, Appl
14	76	100.0	1187	21	US-10-487-901-4336 Sequence 4336, Ap
15	76	100.0	1824	9	US-09-930-329-1 Sequence 1, Appl
16	76	100.0	1824	9	US-09-930-342-1 Sequence 1, Appl
17	76	100.0	1824	24	US-11-004-043-1 Sequence 1, Appl
18	76	100.0	6395	10	US-09-962-527-1 Sequence 1, Appl
19	76	100.0	6395	17	US-10-338-592-2 Sequence 2, Appl
20	76	100.0	6395	19	US-10-828-029-1 Sequence 1, Appl
21	76	100.0	6425	10	US-09-962-527-3 Sequence 3, Appl
22	76	100.0	6425	19	US-10-828-029-3 Sequence 3, Appl
23	76	100.0	6439	10	US-09-962-527-2 Sequence 2, Appl
24	76	100.0	6439	19	US-10-828-029-2 Sequence 2, Appl
25	76	100.0	6446	10	US-09-962-527-5 Sequence 5, Appl
26	76	100.0	6446	19	US-10-828-029-5 Sequence 5, Appl
27	76	100.0	6475	10	US-09-962-527-4 Sequence 4, Appl
28	76	100.0	6475	19	US-10-828-029-4 Sequence 4, Appl
29	76	100.0	7685	9	US-09-949-317-22 Sequence 22, Appl
30	76	100.0	7685	9	US-09-949-317-25 Sequence 25, Appl
31	76	100.0	7685	9	US-09-949-316-22 Sequence 22, Appl
32	76	100.0	7685	9	US-09-949-316-25 Sequence 25, Appl
33	76	100.0	7685	14	US-10-200-051-22 Sequence 22, Appl
34	76	100.0	7685	14	US-10-200-051-25 Sequence 25, Appl
35	76	100.0	7685	22	US-10-624-193-1 Sequence 1, Appl
36	76	100.0	7686	9	US-09-949-317-23 Sequence 23, Appl
37	76	100.0	7686	9	US-09-949-317-26 Sequence 26, Appl
38	76	100.0	7686	9	US-09-949-316-23 Sequence 23, Appl
39	76	100.0	7686	9	US-09-949-316-26 Sequence 26, Appl
40	76	100.0	7686	14	US-10-200-051-23 Sequence 23, Appl
41	76	100.0	7686	14	US-10-200-051-26 Sequence 26, Appl
42	76	100.0	7686	22	US-10-624-193-2 Sequence 2, Appl
43	76	100.0	7687	9	US-09-949-317-24 Sequence 24, Appl
44	76	100.0	7687	9	US-09-949-316-24 Sequence 24, Appl
45	76	100.0	7687	14	US-10-200-051-24 Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-10-487-901-4335/c
; Sequence 4335, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteris
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4335
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-4335

```
Query Match          100.0%; Score 76; DB 21; Length 557;
Best Local Similarity 100.0%; Pred. No. 4.1e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 373 GTGAGACGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 314

Qy 61 GTCCTATGTCGATCA 76
    ||||||||||||||||||
Db 313 GTCCTATGTCGATCA 298

RESULT 2
US-09-775-049-17
; Sequence 17, Application US/09775049
; Publication No. US20030036641A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Fitzmaurice, Wayne P.
; APPLICANT: Lindbo, John A.
; TITLE OF INVENTION: Methods For Homology-Driven Reassembly
; TITLE OF INVENTION: Of Nucleic Acid Sequences
; FILE REFERENCE: P-LG 4412
; CURRENT APPLICATION NUMBER: US/09/775,049
; CURRENT FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 807
; TYPE: DNA
; ORGANISM: tobamovirus
; OTHER INFORMATION: synthetic construct
US-09-775-049-17

Query Match          100.0%; Score 76; DB 10; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GTGAGACGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
    ||||||||||||||||||
Db 601 GTCCTATGTCGATCA 616

RESULT 3
US-10-066-390-10
; Sequence 10, Application US/10066390
; Publication No. US20020146732A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/066,390
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-066-390-10

Query Match          100.0%; Score 76; DB 13; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GTGAGACGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GTGAGACGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
    ||||||||||||||||||
Db 601 GTCCTATGTCGATCA 616

RESULT 4
US-10-206-030-10
; Sequence 10, Application US/10206030
; Publication No. US20020177160A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/206,030
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/066,390
; PRIOR FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-206-030-10

Query Match          100.0%; Score 76; DB 13; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GTGAGACGAGGCGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTTCATGGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
    ||||||||||||||||||
Db 601 GTCCTATGTCGATCA 616

RESULT 5
US-10-211-079-10
; Sequence 10, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vaewhongs, Andrew A.
; APPLICANT: Vojdani, Fakhrieh S.
; APPLICANT: Smith, Mark L.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; TITLE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211,079
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-211-079-10
```

```
; OTHER INFORMATION: synthetic construct
US-10-205-772-10
Query Match      100.0%; Score 76; DB 15; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 60
    |||||||
Db 541 GTGAGACGCGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
    |||||||
Db 601 GTCCTATGTCGATCA 616

RESULT 6
US-10-356-708-9
; Sequence 9, Application US/10356708
; Publication No. US20030157682A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
; FILE REFERENCE: P-LG 10100
; CURRENT APPLICATION NUMBER: US/10/356,708
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,722
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 10/098,155
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 10/211,079
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-10-356-708-9
Query Match      100.0%; Score 76; DB 16; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 60
    |||||||
Db 541 GTGAGACGCGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
    |||||||
Db 601 GTCCTATGTCGATCA 616

RESULT 7
US-10-205-772-10
; Sequence 10, Application US/10205772
; Publication No. US20030186261A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/205,772
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/066,390
; PRIOR FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: synthetic construct
US-10-280-913A-9
Query Match      100.0%; Score 76; DB 16; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 60
    |||||||
Db 541 GTGAGACGCGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
    |||||||
Db 601 GTCCTATGTCGATCA 616

RESULT 8
US-10-280-913A-9
; Sequence 9, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/280,913A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-10-280-913A-9
Query Match      100.0%; Score 76; DB 19; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 60
    |||||||
Db 541 GTGAGACGCGAGGGCCCATGGAACCTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
    |||||||
Db 601 GTCCTATGTCGATCA 616

RESULT 9
US-10-684-134-9
; Sequence 9, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/684,134
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
```

```
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-10-684-134-9

Query Match      100.0%; Score 76; DB 19; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 60
Db 541 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 10
US-10-637-758-9
; Sequence 9, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/637,758
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/10/280,913
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-10-637-758-9

Query Match      100.0%; Score 76; DB 19; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 60
Db 541 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 11
US-10-226-372A-10
; Sequence 10, Application US/10226372A
; Publication No. US20050153283A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method Of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 5414
```

```
; CURRENT APPLICATION NUMBER: US/10/226,372A
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 10/226,372
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-226-372A-10

Query Match      100.0%; Score 76; DB 22; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 60
Db 541 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 12
US-10-624-193-3
; Sequence 3, Application US/10624193
; Publication No. US20050175590A1
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P
; APPLICANT: Pogue, Gregory P
; APPLICANT: Lindbo, John A
; TITLE OF INVENTION: VIRAL EXPRESSION VECTORS
; FILE REFERENCE: 60-016511US
; CURRENT APPLICATION NUMBER: US/10/624,193
; CURRENT FILING DATE: 2003-07-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-624-193-3

Query Match      100.0%; Score 76; DB 22; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 60
Db 541 GTGAGACGCGAGGCCCATCGAACTTACAGAGAAGTCGTTGATGAGTTCATCGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 13
US-10-624-193-4
; Sequence 4, Application US/10624193
; Publication No. US20050175590A1
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P
; APPLICANT: Pogue, Gregory P
; APPLICANT: Lindbo, John A
; TITLE OF INVENTION: VIRAL EXPRESSION VECTORS
; FILE REFERENCE: 60-016511US
; CURRENT APPLICATION NUMBER: US/10/624,193
; CURRENT FILING DATE: 2003-07-21
; NUMBER OF SEQ ID NOS: 6
```


; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-624-193-4

Query Match 100.0%; Score 76; DB 22; Length 807;
Best Local Similarity 100.0%; Pred. No. 4.5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCTTACAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db 541 GTGAGACGCGAGGCCCATCTTACAGAAGTCGTTGATGAGTTTCATGGAAGAT 600

Qy 61 GTCCTATGTCGATCA 76
Db 601 GTCCTATGTCGATCA 616

RESULT 14
US-10-487-901-4336
; Sequence 4336, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; FILE REFERENCE: DOM-0852
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4336
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-4336

Query Match 100.0%; Score 76; DB 21; Length 1187;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCTTACAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db 309 GTGAGACGCGAGGCCCATCTTACAGAAGTCGTTGATGAGTTTCATGGAAGAT 368

Qy 61 GTCCTATGTCGATCA 76
Db 369 GTCCTATGTCGATCA 384

RESULT 15
US-09-930-329-1
; Sequence 1, Application US/09930329
; Patent No. US20020104123A1
; GENERAL INFORMATION:
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: VIRAL AMPLIFICATION OF RECOMBINANT
; TITLE OF INVENTION: MESSENGER RNA IN TRANSGENIC PLANTS
; FILE REFERENCE: 008010103DVUS02
; CURRENT APPLICATION NUMBER: US/09/930,329

; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 09/414,916
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 08/336,724
; PRIOR FILING DATE: 1994-11-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1824
; TYPE: RNA
; ORGANISM: Tobacco Mosaic Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)...(873)
; OTHER INFORMATION: (episomal) peptide encodes fo TMV 30kDa movement
; OTHER INFORMATION: protein (268 residues) and CAT (204 residues).
; NAME/KEY: CDS
; LOCATION: (916)...(1530)
US-09-930-329-1

Query Match 100.0%; Score 76; DB 9; Length 1824;
Best Local Similarity 77.6%; Pred. No. 5.6e-16;
Matches 59; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAGACGCGAGGCCCATCTTACAGAAGTCGTTGATGAGTTTCATGGAAGAT 60
Db 610 GUGAGACGCGAGGCCCAUGGAACUACAGAGAGUCUGUGAUGAGUUCAGGAAGAU 669

Qy 61 GTCCTATGTCGATCA 76
Db 670 GUCCCUAUGUGCAUCA 685

Search completed: September 26, 2005, 05:02:20
Job time : 136 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	197	100.0	1058	6	AX040176	Sequence
2	197	100.0	1058	6	AX795375	Sequence
3	197	100.0	1058	14	SILCP	Satellite
4	194	98.5	198	6	AX040185	Sequence
5	184.2	93.5	1059	6	AX795374	Sequence
6	34.6	17.6	4546	8	COTLEA4A	Sequence
7	34.4	17.5	349980	6	AX344559	Gossypium h
8	34.4	17.5	349980	6	AX344560	Sequence
9	34.2	17.4	845	6	BD272892	Sequence
10	34.2	17.4	845	8	AY036888	Control o
11	34.2	17.4	868	6	BD272888	Brassica
12	34.2	17.4	891	6	BD272894	Control o
13	34.2	17.4	183650	2	AC134713	Control o
14	34.2	17.4	232864	2	AC138224	Rattus no
15	34.2	17.4	235664	2	AC113727	Rattus no
16	34.2	17.4	237432	2	AC113727	Rattus no
17	34.2	17.4	238901	2	AC123182	Rattus no
18	34	17.3	138876	2	AL590067	Rattus no
19	33.6	17.1	14376	1	AE013049	Homo sapi
20	33.6	17.1	14376	1	AE013049	Thermoana

```
Db 181 AACCAACCGTAATCG 197
|||||
RESULT 2
AX795375 1058 bp mRNA linear PAT 04-OCT-2003
LOCUS Sequence 2 from Patent WO03052108.
ACCESSION AX795375
VERSION AX795375.1 GI:37516048
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Metzlaiff,M.H., Gosselle,V.M., Meulewaeter,F. and Fache,I.C.
TITLE Improved methods and means for delivering inhibitory rna to plants
and applications thereof
JOURNAL Patent: WO 03052108-A 2 26-JUN-2003;
Bayer BioScience N.V. (BE)
FEATURES
source Location/Qualifiers
1..1058
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/notice="cdna sequence of the genome of satellite tobacco
mosaic virus"
ORIGIN
Query Match 100.0%; Score 197; DB 6; Length 1058;
Best Local Similarity 100.0%; Pred. No. 3e-47; Indels 0; Gaps 0;
Matches 197; Conservative 0; Mismatches 0;
Qy 1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTTATGCTGTTG 60
|||||
Db 1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTTATGCTGTTG 60
|||||
Qy 61 GGGGACATAGGGGGAACAATATTCCTTCTTACAAGAGGCTTCAGTGCCTCAATATT 120
|||||
Db 61 GGGGACATAGGGGGAACAATATTCCTTCTTACAAGAGGCTTCAGTGCCTCAATATT 120
|||||
Qy 121 ACTTGGCGCCCAATTTGGGTTTCAGTTCGTTTCCAGCTATCGGGAGAGGTAAGTTA 180
|||||
Db 121 ACTTGGCGCCCAATTTGGGTTTCAGTTCGTTTCCAGCTATCGGGAGAGGTAAGTTA 180
|||||
Qy 181 AACCAACCGTAATCG 197
|||||
Db 181 AACCAACCGTAATCG 197
|||||
RESULT 3
SLCPC 1058 bp ss-RNA linear VRL 16-AUG-1994
LOCUS Satellite tobacco mosaic virus coat protein RNA, complete cds.
ACCESSION M25782
VERSION M25782.1 GI:530201
KEYWORDS coat protein.
SOURCE Satellite tobacco mosaic virus
ORGANISM Satellite tobacco mosaic virus
REFERENCE 1
AUTHORS Mirkov,T.E., Mathews,D.M., Du Plessis,D.H. and Dodds,J.A.
TITLE Nucleotide sequence and translation of satellite tobacco mosaic
virus RNA
JOURNAL Virology 170 (1), 139-146 (1989)
MEDLINE 89243170
PUBMED 2718378
COMMENT On Aug 17, 1994 this sequence version replaced gi:341395.
Original source text: Satellite tobacco mosaic virus RNA.
FEATURES
source Location/Qualifiers
1..1058
/organism="Satellite tobacco mosaic virus"
/mol_type="genomic RNA"
```

```
CDS
/db_xref="taxon:12254"
52..228
/notes="ORF"
/codon_start=1
/protein_id="AAA47784.1"
/db_xref="GI:530202"
/tranlations="MULLGDIGGKHIAFFYKRPSSVAIITWRPILGFSCCFQLWGEVRLN
QTVNRVVTIRMLL"
162..641
/codon_start=1
/product="coat protein"
/protein_id="AAA47785.1"
/db_xref="GI:530203"
/tranlations="MGRGKVKPNRKSTGDSNVVTMIRAGSYPKVNPPTWVRAIPPE
VSQSGIAFKVPVGSFNSFRVTVMSVRVATQITPPVNEFSFVRLKELFKT
GDSTEEFEGRASININRASVGYRIPNLRQNTVAADNVCEVRSNCRQVALVISCCFN"
ORIGIN
Query Match 100.0%; Score 197; DB 14; Length 1058;
Best Local Similarity 100.0%; Pred. No. 3e-47; Indels 0; Gaps 0;
Matches 197; Conservative 0; Mismatches 0;
Qy 1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTTATGCTGTTG 60
|||||
Db 1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTTATGCTGTTG 60
|||||
Qy 61 GGGGACATAGGGGGAACAATATTCCTTCTTACAAGAGGCTTCAGTGCCTCAATATT 120
|||||
Db 61 GGGGACATAGGGGGAACAATATTCCTTCTTACAAGAGGCTTCAGTGCCTCAATATT 120
|||||
Qy 121 ACTTGGCGCCCAATTTGGGTTTCAGTTCGTTTCCAGCTATCGGGAGAGGTAAGTTA 180
|||||
Db 121 ACTTGGCGCCCAATTTGGGTTTCAGTTCGTTTCCAGCTATCGGGAGAGGTAAGTTA 180
|||||
Qy 181 AACCAACCGTAATCG 197
|||||
Db 181 AACCAACCGTAATCG 197
|||||
RESULT 4
AX040185 198 bp DNA linear PAT 18-NOV-2000
LOCUS Sequence 13 from Patent WO0063397.
ACCESSION AX040185
VERSION AX040185.1 GI:11230135
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Meulewaeter,F., Cornelisse, Jacobs,J., van Eldik,G. and Metzlaiff,M.
TITLE Methods and means for delivering inhibitory rna to plants and appli
cations thereof
JOURNAL Patent: WO 0063397-A 13 26-OCT-2000;
Aventis CropScience N.V. (BE)
FEATURES
source Location/Qualifiers
1..198
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notice="cdna copy of STMV leader region"
ORIGIN
Query Match 98.5%; Score 194; DB 6; Length 198;
Best Local Similarity 100.0%; Pred. No. 3e-46;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 AAATTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTTATGCTGTTGGG 63
|||||
Db 5 AAATTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTTATGCTGTTGGG 64
|||||
Qy 64 GACATAGGGGGAACAATATTCCTTCTTACAAGAGGCTTCAGTGCCTCAATATTACT 123
|||||
```

Db 65 GACATAGGGGAAACATATTGCGCTTCTTCTACAGAGCGCTTCAGTCGCCATTAATTAAT 124

Qy 124 TGGCGCCCAATTTGGGTTTCAGTTGCTGTTTCCAGCTATGGGAGAGGTAAGGTTAAAC 183

Db 125 TGGCGCCCAATTTGGGTTTCAGTTGCTGTTTCCAGCTATGGGAGAGGTAAGGTTAAAC 184

Qy 184 CAACACCGTAATTCG 197

Db 185 CAACACCGTAATTCG 198

RESULT 5

AX795374

LOCUS AX795374 1059 bp mRNA linear PAT 04-OCT-2003

DEFINITION Sequence 1 from Patent WO03052108.

ACCESSION AX795374

VERSION AX795374.1 GI:37516047

KEYWORDS

SOURCE

ORGANISM

synthetic construct

other sequences; artificial sequences.

REFERENCE

AUTHORS Metzloff,M.H., Gossele,V.M., Meulewaeter,F. and Fache,I.C.

TITLE Improved methods and means for delivering inhibitory rna to plants and applications thereof

JOURNAL Patent: WO 03052108-A 1 26-JUN-2003;

Bayer BioScience N.V. (BE)

FEATURES

source

1. 1059

/location/Qualifiers

/organism="synthetic construct"

/mol_type="mRNA"

/db_xref="taxon:32630"

/note="cDNA sequence of the genome of satellite tobacco mosaic virus cloned into pSMV-1"

ORIGIN

Query Match 93.5%; Score 184.2; DB 6; Length 1059;

Best Local Similarity 95.9%; Pred. No. 1.8e-43;

Matches 189; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 AGTAACTTACCAATCAATCAAGACCTTACCAACAGAGCTGCTGGTCAATTTATGCTGTG 60

Db 1 AGTAACTTACCAATCAAGACCTTACCAACAGAGCTGCTGGTCAATTTATGCTGTG 60

Qy 61 GGGGACATAGGGGAAACATATTGCGCTTCTTCTACAGAGCGCTTCAGTCGCCATAATT 120

Db 61 GGGGACATAGGGGAAACATATTGCGCTTCTTCTACAGAGCGCTTCAGTCGCCATAATT 120

Qy 121 ACTTGGGCGCCAAATTTGGGTTTCAGTTGCTGTTTCCAGCTATGGGAGAGGTAAGGTTA 180

Db 121 ACTTGGGCGCCAAATTTGGGTTTCAGTTGCTGTTTCCAGCTATGGGAGAGGTAAGGTTA 180

Qy 181 AACCAACCGTAATTCG 197

Db 181 AACCAACCGTAATTCG 197

RESULT 6

COTLEA4A

LOCUS Gossypium hirsutum Le4-A gene, complete CDS. 4546 bp DNA linear PLN 27-APR-1993

DEFINITION

ACCESSION M73752

VERSION M73752.1 GI:167336

KEYWORDS Le4-A protein.

SOURCE Gossypium hirsutum (upland cotton)

ORGANISM Gossypium hirsutum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 4546)

Galau,G.A., Wang,H.Y.-C. and Hughes,D.W.

Sequence of Le4 (D19) from Gossypium hirsutum: The A-genome allele, the D-genome allele and the correct sequence

JOURNAL

COMMENT

of their proteins

Unpublished (1991)

Original source text: Gossypium hirsutum embryo before endoreduplication of DNA cotyledon DNA.

The sequence of a cDNA encoded by the D allele is available under accession number M73751 and is used to define introns, exons, poly(A) site, etc.

Qualifiers

1. 4546

/organism="Gossypium hirsutum"

/mol_type="genomic DNA"

/db_xref="taxon:3635"

/tissue_type="cotyledon"

/dev_stage="embryo before endoreduplication of DNA"

2176. 2181

/note="putative"

join(<2226. 2418,2520. 3041)

/genes="Lea4-A"

join(<2226. 2418,2520. 3041)

/genes="Lea4-A"

/product="Lea4-A protein gene"

/function="possible desiccation protectant"

/evidence="experimental"

<2226. 2418

/genes="Lea4-A"

/product="Lea4-A protein"

/standard_name="Lea4-A"

/number=1

/function="possible desiccation protectant"

/evidence="experimental"

join(2274. 2418,2520. 2683)

/genes="Lea4-A"

/function="possible desiccation protectant"

/standard_name="Lea4-A protein gene"

/codon_start=1

/evidence="experimental"

/protein_id="AAA33060.1"

/db_xref="GI:167337"

/translation="MASEQYQAMRNAPQEEKELDARAKQGETVPGGTGKSLDAQI NLAEGRHKGGETRKQQLGTGTYQMGRKGLSNSDMSGGERAADEGVTIDSKPRTKN"

2419. 2519

/genes="Lea4-A"

2520. 3041

/genes="Lea4-A"

/number=2

/evidence="experimental"

2859. 4028

/note="putative"

/rpt_type=tandem

3041

/gene="Lea4-A"

ORIGIN

Query Match 17.6%; Score 34.6; DB 8; Length 4546;

Best Local Similarity 61.8%; Pred. No. 17;

Matches 55; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 4 AAATCTACCAATCAAGACCTAACCAACAGGACTGCTGCTCATTTATGCTGTTGGG 63

Db 2026 AAGTTTACCAATCAAGACCTAACCAACAGGACTGCTGCTCATTTATGCTGTTGGT 2085

Qy 64 GACATAGGGGAAACATATTGCTTCTT 92

Db 2086 TGCTTACCTGATCACAATTTGTTCTT 2114

RESULT 7

AX344559

LOCUS AX344559 34980 bp DNA linear PAT 01-FEB-2002

DEFINITION Sequence 10 from Patent WO0200932.

ACCESSION AX344559

VERSION AX344559.1 GI:18492445


```

/db_xref="taxon:3708"

ORIGIN
Query Match      17.4%; Score 34.2; DB 6; Length 845;
Best Local Similarity 67.6%; Pred. No. 30;
Matches 48; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 88 TTCTTCTACAGAGCCCTTCAGTCGCCCAATAATTACTTGGCGCCCAATTTTGGGTTTCAGT 147
DB 81 TTCTTCTCTCCCATGGCTTCTGTCTCCCAACAGTGCTTTGGCGCCCTAAATTGATCTTGAGG 22

QY 148 TGCTGTTTCCA 158
DB 21 TCCGGTTTCCA 11

RESULT 10
LOCUS AY036888/c 845 bp mRNA linear PLN 19-DEC-2001
DEFINITION Brassica napus MADS-box protein (FLC1) mRNA, complete cds.
ACCESSION AY036888
VERSION AY036888.1 GI:17933449
KEYWORDS
SOURCE
ORGANISM
Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
AUTHORS Tadege,M., Sheldon,C.C., Helliwell,C.A., Stoutjesdijk,P.,
Dennis,E.S. and Peacock,W.J.
TITLE Control of flowering time by FLC orthologues in Brassica napus
JOURNAL Plant J. 28 (5), 545-553 (2001)
MEDLINE 21839898
PUBMED 11849594
REFERENCE
AUTHORS Helliwell,C.A. and Tadege,M.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2001) Plant Industry, CSIRO, Black Mountain,
Canberra, ACT 2601, Australia
FEATURES
source
1..845
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Colombus"
/db_xref="taxon:3708"
1..845
/gene="FLC1"
68..661
/gene="FLC1"
/note="flowering locus C ortholog"
/codon_start=1
/product="MADS-box protein"
/protein_id="AAK70215.1"
/db_xref="GI:17933450"
/translation="WGRKKLRIKRIENKSRQVTFSKRRNGLIEKAROLSVLCDASVA
LLVVSAGKLYFSFGNDLVKILDRYGHODDLKALDRQSKALDCSHHLLLELVES
KLESNDVNVSVGSLVQLVEHLENALSVTRARKTELMLKLVLENLKEKELLEEENHVL
ASQMEKSNLVRADNMDVSPGQISDINLPLPLLN"

gene
CDS

ORIGIN
Query Match      17.4%; Score 34.2; DB 8; Length 845;
Best Local Similarity 67.6%; Pred. No. 30;
Matches 48; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 88 TTCTTCTACAGAGCCCTTCAGTCGCCCAATAATTACTTGGCGCCCAATTTTGGGTTTCAGT 147
DB 81 TTCTTCTCTCCCATGGCTTCTGTCTCCCAACAGTGCTTTGGCGCCCTAAATTGATCTTGAGG 22

QY 148 TGCTGTTTCCA 158
DB 21 TCCGGTTTCCA 11

RESULT 11
LOCUS BD272888/c 868 bp DNA linear PAT 17-JUL-2003
DEFINITION Control of flowering.
ACCESSION BD272888
VERSION BD272888.1 GI:33082656
KEYWORDS JP 2002532069-A/7.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
AUTHORS Burn,J.E., Peacock,W.J., Dennis,E.S., Sheldon,C.C., Perez,P.,
Helliwell,C.A. and Rouse,D.T.
TITLE Control of flowering
JOURNAL Patent: JP 2002532069-A 7 02-OCT-2002;
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION,
PASCUAL PEREZ
COMMENT OS Brassica napus (rape)
PN JP 2002532069-A/7
PD 02-OCT-2002
PF 02-DEC-1999 JP 2000585411
PR 03-DEC-1998 AU PP 7469,22-JAN-1999 US 60/116928 PI
JOANNE ELIZABETH BURN,WILLIAM JAMES PEACOCK,ELIZABETH PI
SALISBURY DENNIS,
PI CANDICE CLAIRE SHELDON,PASCUAL PEREZ,CHRISTOPHER ANDREW PI
HELLIWELL,
PI DEAN THOMAS ROUSE
PC C12N15/09,A01H5/00,C07K14/415,C07K16/16,C12N5/10,C12Q1/68,PC
G01N33/48,
PC G01N33/53,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC Control
of flowering
FH Key Location/Qualifiers
FT source 1..868
FT Location/Qualifiers
1..868
/organism="Brassica napus (rape)".
/organism="Brassica napus"
/mol_type="genomic DNA"
/db_xref="taxon:3708"

ORIGIN
Query Match      17.4%; Score 34.2; DB 6; Length 868;
Best Local Similarity 67.6%; Pred. No. 30;
Matches 48; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 88 TTCTTCTACAGAGCCCTTCAGTCGCCCAATAATTACTTGGCGCCCAATTTTGGGTTTCAGT 147
DB 93 TTCTTCTCTCCCATGGCTTCTGTCTCCCAACAGTGCTTTGGCGCCCTAAATTGATCTTGAGG 34

QY 148 TGCTGTTTCCA 158
DB 33 TCCGGTTTCCA 23

RESULT 12
LOCUS BD272894/c 891 bp DNA linear PAT 17-JUL-2003
DEFINITION Control of flowering.
ACCESSION BD272894
VERSION BD272894.1 GI:33082662
KEYWORDS JP 2002532069-A/13.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
AUTHORS Burn,J.E., Peacock,W.J., Dennis,E.S., Sheldon,C.C., Perez,P.,
Helliwell,C.A. and Rouse,D.T.
TITLE Control of flowering

```

JOURNAL Patent: JP 2002532069-A 13 02-OCT-2002;
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION,
PASCUAL PEREZ

COMMENT OS Brassica napus (rape)
PN JP 2002532069-A/13
PD 02-OCT-2002
PR 02-DEC-1999 JP 2000585411
PP 7469, 22-JAN-1999 US 60/116928 PI
JOANNE ELIZABETH BURN, WILLIAM JAMES PEACOCK, ELIZABETH PI
SALISBURY DENNIS.
PI CANDICE CLAIRE SHELTON, PASCUAL PEREZ, CHRISTOPHER ANDREW PI
HELLIWELL,
PI DEAN THOMAS ROUSE
PC C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12Q1/68, PC
GOIN33/48,
PC GOIN33/53, GOIN33/53, GOIN33/566, C12N15/00, C12N5/00 CC Control
of flowering
FH Key Location/Qualifiers
FT source 1. .891
FT Location/Qualifiers
/organism='Brassica napus (rape)'.
1. .891
/organism='Brassica napus'
/mol_type='genomic DNA'
/db_xref='taxon:3708'

ORIGIN Query Match 17.4%; Score 34.2; DB 6; Length 891;
Best Local Similarity 67.6%; Pred. No. 30;
Matches 48; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 88 TTCTCTACAGAGCCCTTCAGTCGCATATTACTTCGCGCCCAATTTTGGTTTCAGT 147
|||||
DB 80 TTCTTCTCCCAAGGCTTCGTCTCTCCACAGTGCTTTGCGCCCTAATTGATCTTGAGG 21
|||||

QY 148 TGCTGTCTTCCA 158
|||||
DB 20 TCCGGTTCCA 10
|||||

RESULT 13
AC134713

LOCUS 183650 bp DNA linear HTG 20-NOV-2002
Rattus norvegicus clone CH230-453115, WORKING DRAFT SEQUENCE, 2
unordered pieces.

AC134713
AC134713.2 GI:25139142
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 183650)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alebrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,O.K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J.J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flag,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Manhum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokeleneh,O., Okwuon,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,W., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 183650)
Rat Genome Sequencing Consortium.
Submitted (30-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 183650)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23346392.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBSF
Center clone name: CH230-453115
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 164086 bases at least Q40
Consensus quality: 167423 bases at least Q30
Consensus quality: 168939 bases at least Q20
Estimated insert size: 164689; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 134648: contig of 134648 bp in length
 * 134649 134748: gap of unknown length
 * 134749 183650: contig of 48902 bp in length.

FEATURES
 source
 Location/Qualifiers
 1..183650
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-453115"
 1..1028
 /note="wgs end extension
 clone_end:Sp6"
 complement(5130..6053)
 /note="clone boundary
 clone_end:Sp6
 site:
 end_sequence:BZ146734"
 70736..134356
 /note="clone boundary
 clone_end:T7"
 site:
 end_sequence:BZ146733"
 134749..136628
 /note="wgs end extension
 clone_end:T7"

ORIGIN
 Query Match 17.4%; Score 34.2; DB 2; Length 183650;
 Best Local Similarity 56.8%; Pred. No. 12;
 Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 5 AACTTACCAATCAAGACCTTAACCAACAGACTGCTGTCATTATGCTGTGGGG 64
 Db |||||
 41701 AACTGACAAATCAACAAATCAACACAGAAAGTGTCTGTCCTCTATGCGG 41760
 |||||

QY 65 ACATAGGGGAAACATATTCCTCTCAAGAGCCCTTCAGTCGCCA 115
 Db |||||
 41761 TCCAGGCAACACCATCACTTCAAGAAAGAACTTAAGTTTCCA 41811
 |||||

RESULT 14
 AC128224/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-497M8, WORKING DRAFT SEQUENCE, 5
 unorderd pieces.
 ACCESSION AC128224
 VERSION AC128224.4 GI:25188214
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 232864)
 Murny,D.Marie., Metzker,M.Lee., Abranzon,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,I., Garza,M.,
 Guebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Loresuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,B.,
 Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokeleneh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
 Popper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P.,
 Williams,G., Wilson,R., Wlezyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 232864)
 Worley,K.C.
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 232864)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (23-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 23, 2002 this sequence version replaced gi:23269495.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information

COMMENT

```
Center project name: GZLS
Center clone name: CH230-497M8
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329
Consensus quality: 210769 bases at least Q40
Consensus quality: 213407 bases at least Q30
Consensus quality: 214662 bases at least Q20
Estimated insert size: 210905; sum-of-ctnigs estimation
Quality coverage: 7x in Q20 bases; sum-of-ctnigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
      (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 5 ctnigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the ctnigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  *
  * 1 10076: ctnig of 10076 bp in length
  * 10077 10176: gap of unknown length
  * 10177 13489: ctnig of 3313 bp in length
  * 13490 13589: gap of unknown length
  * 13590 219491: ctnig of 205902 bp in length
  * 219492 219591: gap of unknown length
  * 219592 231352: ctnig of 11761 bp in length
  * 231353 231452: gap of unknown length
  * 231453 232864: ctnig of 1412 bp in length.
  * 231453 Location/Qualifiers
FEATURES
    source
        1..232864
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-497M8"
            10177..11235
                /note="wgs_ctnig"
            13590..14542
                /note="wgs_ctnig"
            219592..221536
                /note="wgs_ctnig"
ORIGIN
Query Match          17.4%; Score 34.2; DB 2; Length 232864;
Best Local Similarity 50.3%; Pred. No. 12;
Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 27 ACCAACAGGACTGTCTGCTCATTTATCTGTGGGACATAGGGGAAAAACATATTGC 86
Db 150370 AACTCGGAATTACTGTAGGGTTTTTGGAGACAGACAGATCCTGGGTATCTTTCT 150311
QY 87 CTTCTTCTACAGAGCCCTTCAGTCCCATTAATTAATCTTGGCGCCCAATTTGGGTTTCAG 146
Db 150310 CTTCTCTCAGAGCCCTCTGCTCTCAGAGAAATCTTTGAGCGGAAGTCAGAGAAGGC 150251
QY 147 TTGCTGTTTCCAGCTATCGGAGAGGTAGGTTAAACCAACCCGTAA 193
Db 150250 TTCCAAATACCAGCTTTGGGCATAGCATACTTTATAGTACTCTCTTA 150204

RESULT 15
AC113727/c
LOCUS AC113727 235664 bp DNA linear HTG 19-NOV-2002
DEFINITION Rattus norvegicus clone CH230-4709, *** SEQUENCING IN PROGRESS ***,
3 unordered pieces.
ACCESSION AC113727
VERSION AC113727-5 GI:25072682
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
```

REFERENCE

AUTHORS

1 (bases 1 to 235664)

Muzny, D., Marie., Metzker, M., Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyatebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, R., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.T., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulsegod, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwankweme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 235664)

Worley, K.C.

Direct Submission

Submitted (05-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 235664)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265862.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence

REFERENCE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

JOURNAL

AUTHORS

TITLE

JOURNAL

COMMENT

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSDG
Center clone name: CH230-4709
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 222163 bases at least Q40
Consensus quality: 224226 bases at least Q30
Consensus quality: 225338 bases at least Q20
Estimated insert size: 224820; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 227304: contig of 227304 bp in length
* 227305 227404: gap of unknown length
* 227405 228430: contig of 1026 bp in length
* 228431 228530: gap of unknown length
* 228531 235664: contig of 7134 bp in length.

FEATURES

source	1..235664
	Location/Qualifiers
	1..235664
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-4709"
misc_feature	1..1010
	/note="wgs contig"
misc_feature	1131..2787
	/note="wgs contig"
misc_feature	107373..108409
	/note="wgs contig"
misc_feature	220040..220922
	/note="clone boundary"
	clone_end:Sp6
	site:
	end_sequence:BH361801"
misc_feature	225054..227304
	/note="wgs end extension"
	clone_end:Sp6"

ORIGIN

Query Match	17.4%	Score 34.2;	DB 2;	Length 235664;
Best Local Similarity	50.3%	Pred.No.12;		
Matches	84;	Conservative 0;	Mismatches 83;	Indels 0;
			Gaps 0;	
Qy	27	ACCAACAGGACTGTCTGCTCATTTCTGTGGGACATAGGGGAAACATATTGC	86	
Db	232651	AACGGGGAATTACTGTAGGGTTTTTGGACAGCAGAGATCTGGGTATCTCTTCT	232592	
Qy	87	CTTCTTTTCAAGAGGCTTCAGTCGCCAATAATTCTTGGCCCCCAATTTTGGGTTTCAG	146	
Db	232591	CTTCTCTAGAAGCTCTCTTGTCTCAGAGAACATTTCTTTGAGCGAGTCAGAGAAGGC	232532	
Qy	147	TTCCTGTTTCCAGCTATCGGAGAGGTAAGGTTAAACCAACCGTAA	193	
Db	232531	TTCCAATTACCAGCTTTGGGCATAGCATCTTTATAGCTACCTCTTA	232485	

Search completed: September 26, 2005, 03:05:11
Job time : 1390.47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 23:45:48 ; Search time 247.69 Seconds
(without alignments)
4708.259 Million cell updates/sec

Title: US-09-551-494-4_COPY_1_197
Perfect score: 197
Sequence: 1 agtaaacattccaatcaaaa.....ttaaaccaaacccgtaaatcg 197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	197	100.0	1058	3 AAC62371	Aac62371 cDNA sequ
2	197	100.0	1058	9 ACC85000	Acc85000 STMV geno
3	197	100.0	1058	12 ADOS8866	Ado58866 Satellite
4	194	98.5	198	3 AAC62380	Aac62380 Leader re
5	194	98.5	1059	2 AAQ06316	Aaq06316 Satellite
6	184.2	93.5	1059	9 ACC84999	Acc84999 STMV geno
7	48.8	24.8	62	2 AAQ06899	Aaq06899 Probe to
8	34.2	17.4	845	3 AAA29039	Aaa29039 B. napus
9	34.2	17.4	868	3 AAA29035	Aaa29035 B. napus
10	34.2	17.4	891	3 AAA29041	Aaa29041 B. napus
11	33.2	16.9	198285	6 ABK84699	Abk84699 Human cDN
12	33.2	16.9	198285	6 ABN97319	Abn97319 Gene #381
13	33.2	16.9	198285	13 ADR52987	Adr52987 Drug ther
14	30.6	15.5	144723	11 ACN44898	Acn44898 Human gen
15	30.4	15.4	819	11 ACN84908	Acn84908 Breast ca
16	30.4	15.4	1196	11 ACN84320	Acn84320 Breast ca
17	30	15.2	414	10 ACF71385	Acf71385 Photorhab
18	30	15.2	43576	3 AAA68609	Aaa68609 Bacteriop
19	30	15.2	110000	10 ACP67367_45	Continuation (46 o
20	30	15.2	110000	10 ACP65388_02	Continuation (3 of

21	30	15.2	110000	12 ADJ25985_02	Continuation (3 of
22	30	15.2	110000	12 ADJ25985_03	Continuation (4 of
C 23	30	15.2	249487	6 ABN85733	Abn85733 Mouse gen
C 24	29.6	15.0	622	5 ABV52317	Abv52317 Human pro
C 25	29.4	14.9	354	8 ACF73082	Acf73082 Staphyloc
26	29.4	14.9	518	5 ADI75534	Adi75534 Human ova
27	29.4	14.9	518	5 ADI69187	Adi69187 Human ova
28	29.4	14.9	1989	10 ADH84392	Adh84392 Enterococ
29	29.4	14.9	8726	2 AAX13032	Aax13032 Enterococ
30	29.4	14.9	8726	6 ABS98827	Abes98827 Enterococ
31	29.4	14.9	10401	2 AAV74460	Aav74460 Staphyloc
32	29.4	14.9	14416	4 ABL14206	Abi14206 Drosophil
33	29.4	14.9	31271	4 ABL26680	Abi26680 Drosophil
34	29.4	14.9	110000	6 ABX08336_02	Continuation (3 of
35	29.4	14.9	110000	6 ABX08336_03	Continuation (4 of
36	29.4	14.9	110000	12 ADN97989_02	Continuation (3 of
37	29.4	14.9	110000	12 ADN97989_03	Continuation (4 of
38	29.4	14.9	110000	12 ADOS0281_02	Continuation (3 of
39	29.4	14.9	110000	12 ADOS0281_03	Continuation (4 of
C 40	29.2	14.8	158811	12 ADQ19901	Adq19901 Human sof
41	29	14.7	13788	5 ABA19999	Abal9999 Human ner
C 42	29	14.7	209484	11 ACN44126	Acn44126 Human gen
C 43	28.8	14.6	399	4 AAI82962	Aai82962 Human pol
C 44	28.8	14.6	862	10 ADK58242	Adk58242 Plant DNA
45	28.8	14.6	862	10 ADK54149	Adk54149 Plant DNA

ALIGNMENTS

RESULT 1
AAC62371
ID AAC62371 standard; DNA; 1058 BP.

XX AAC62371;

DT 19-MAR-2001 (first entry)

XX cDNA sequence of the genome of satellite tobacco mosaic virus.

XX Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
KW Gossypium; Triticum; Arabidopsis; Pecunia; herbicide; transgenic plant;
KW tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus; ss.

OS Satellite tobacco mosaic virus.

PN W0200063397-A2.

XX 26-OCT-2000.

XX 17-APR-2000; 2000WO-EP003521.

XX 20-APR-1999; 99US-00294022.

XX (AVET) AVENTIS CROPS SCIENCE NV.

XX Meulewaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaiff M;

XX WPI; 2000-687182/67.

XX Identifying and isolating genes involved in determining the trait or
phenotype of plant species, by infecting plants with gene silencing
constructs targeted to the gene, and identifying plants with altered
traits.

XX Example 1; Page 53; 64pp; English.

XX The specification describes a method for isolating genes that determine a
trait or phenotype of a plant species. The method comprises identifying a
set of nucleic acids of genes correlated with the trait, creating a
library of gene silencing constructs in a viral RNA vector, targeting the
gene silencing constructs to the nucleic acid set, infecting a collection
of individual plants with these, identifying plants with altered traits

or phenotype, and isolating genes of the invention. The method is useful for isolating genes involved in the determination of trait or a phenotype of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypium, Triticum, Arabidopsis or Petunia. The method is also useful for modulating the expression of selected nucleic acid sequences and for validating the function of a nucleic acid sequence whose expression is correlated with the presence or absence of a specific trait in plants, but with otherwise unknown function. The method is also useful for developing agronomically useful products such herbicides or transgenic plants. The present sequence represents the cDNA sequence of the genome of satellite tobacco mosaic virus (STMV). The sequence was used to construct a plasmid vector for the synthesis of an infective hybrid satellite tobacco mosaic virus (STMV)/satellite tobacco necrosis virus (STNV) helper virus RNA. This helper virus is used in the method of the invention

Sequence 1058 BP; 275 A; 227 C; 258 G; 298 T; 0 U; 0 Other;
Query Match 100.0%; Score 197; DB 3; Length 1058;
Best Local Similarity 100.0%; Pred. No. 3.7e-58;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTATGCTGTG 60
1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTATGCTGTG 60
61 GGGACATAGGGGGAACAATATGCTTCTTCTACAAGAGGCTTCAGTCGCCATAATT 120
61 GGGACATAGGGGGAACAATATGCTTCTTCTACAAGAGGCTTCAGTCGCCATAATT 120
121 ACTTGGCGCCCAATTTGGGTTTCAGTTGCTGTTTCCAGCTATCGGAGAGGTAAGGTTA 180
121 ACTTGGCGCCCAATTTGGGTTTCAGTTGCTGTTTCCAGCTATCGGAGAGGTAAGGTTA 180
181 AACCAACCGTAATCG 197
181 AACCAACCGTAATCG 197

RESULT 2

ACC85000
ID ACC85000 standard; cDNA; 1058 BP.

ACC85000;

13-OCT-2003 (first entry)

STMV genome nucleotide sequence.

Inhibitory RNA; viral RNA vector; coat protein; STMV; gene; ss.

Tobacco mosaic virus.

WO2003052108-A2.

26-JUN-2003.

05-DEC-2002; 2002MO-EP013964.

18-DEC-2001; 2001US-0340488P.

(FARB) BAYER BIOSCIENCE NV.

Metzlaff MH, Gossele VML, Meulewaeter F, Fache ICA;

WPI; 2003-523529/49.

Introducing inhibitory RNA into a plant cell comprises providing a viral RNA vector derived from a satellite RNA virus that encodes a coat protein, and infecting a plant with the viral RNA vector and a corresponding helper virus.

Example 1; Page 75-76; 86pp; English.

The invention relates to introducing inhibitory RNA into a plant cell. The method involves providing a viral RNA vector derived from a satellite RNA virus having a sequence that encodes a coat protein, and infecting a plant with the viral RNA vector and a corresponding helper virus. The methods and viral RNA vectors are useful in introducing inhibitory RNA into plant cells. These may be used to determine or validate the function of isolated nucleic acid sequences in plants. The present sequence represents the nucleotide sequence of the genome of satellite tobacco mosaic virus (STMV) (GenBank Accession No. M25782)

Sequence 1058 BP; 275 A; 227 C; 258 G; 298 T; 0 U; 0 Other;

Query Match 100.0%; Score 197; DB 9; Length 1058;

Best Local Similarity 100.0%; Pred. No. 3.7e-58;

Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTATGCTGTG 60
1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTCGTGTCATTATGCTGTG 60
61 GGGACATAGGGGGAACAATATGCTTCTTCTACAAGAGGCTTCAGTCGCCATAATT 120
61 GGGACATAGGGGGAACAATATGCTTCTTCTACAAGAGGCTTCAGTCGCCATAATT 120
121 ACTTGGCGCCCAATTTGGGTTTCAGTTGCTGTTTCCAGCTATCGGAGAGGTAAGGTTA 180
121 ACTTGGCGCCCAATTTGGGTTTCAGTTGCTGTTTCCAGCTATCGGAGAGGTAAGGTTA 180
181 AACCAACCGTAATCG 197
181 AACCAACCGTAATCG 197

RESULT 3

AD058866

ID AD058866 standard; DNA; 1058 BP.

AD058866;

26-AUG-2004 (first entry)

Satellite tobacco mosaic virus (STMV) genomic sequence.

Multi-component nucleic acid detector; nucleic acid captor;

tail sequence; loop sequence; detector probe; nucleic acid target;

Satellite tobacco mosaic virus; STMV; ds.

Satellite tobacco mosaic virus.

US2004110141-A1.

10-JUN-2004.

10-DEC-2002; 2002US-00315826.

10-DEC-2002; 2002US-00315826.

(UYAL-) UNIV ALABAMA.

Pusey ML, Chittur KK, Dowell JJ, Ng JD;

WPI; 2004-440330/41.

Multi-component nucleic acid detector comprises immobilized nucleic acid captor sequence with hairpin structure, loop sequence complementary to target sequence, and separate detector labeled probe complementary to sequence on captor.

Example; SEQ ID NO 4; 20pp; English.

The present invention relates to a multi-component nucleic acid detector that has a nucleic acid captor sequence bound to a substance at its 5',

terminus, and comprises, from the 5' terminus, a first tail sequence, an optional spacer, a loop sequence complementary to a sequence within the target, an optional second spacer and a second tail sequence complementary to the first, and a detector probe complementary to second tail of capitor and having at least one attached label. Also disclosed is a method for detecting the presence of a nucleic acid target within a biological sample. In the nucleic acid detector, the tail sequences comprise a series of four repeat tetramers. The tail sequences are between 8-20 residues in length, the loop sequence is between 10-40 residues in length, and the optional spacer sequences are from 0-10 nucleotides in length. The capitor and detector probe comprise DNA, RNA, PNA, or its combinations. The detector probe is labelled with a molecule chosen from fluorophores and radioisotopes. The label molecule is chosen from fluorescein, Lucifer yellow, erythrosine, tetramethylrhodamine and coumarin. The nucleic acid detector is useful in detecting the presence of nucleic acid targets within a biological sample. The detector provides for detection of nucleic acid targets with less interference from background noise because detector probes are completely washed from the capitors when no targets are present. This overcomes many of the complications in previous detection methods that relied upon the orientation of labelled probes rather than the presence or absence of the probes. The detector also provides ease in detection of multiple targets when the capitors are disposed as separate clusters upon the surface of a common substrate therefore allowing the simultaneous detection of multiple targets within a common sample. The present sequence represents Satellite tobacco mosaic virus (STMV) genomic sequence. The sequence is used in the exemplification of the method of the invention.

Query Match 100.0%; Score 197; DB 12; Length 1058;
Best Local Similarity 100.0%; Pred. No. 3.7e-58;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGTAACTTACCAATCAAAAGACCTTAACCAAGAGCTGTGCTGCTCATTTATGCTGTG 60
1 AGTAACTTACCAATCAAAAGACCTTAACCAAGAGCTGTGCTGCTCATTTATGCTGTG 60
61 GGGGACATAGGGGAAAACATATGCTCTTCTTACAAGAGGCTTTCAGTGCCTCATATT 120
61 GGGGACATAGGGGAAAACATATGCTCTTCTTACAAGAGGCTTTCAGTGCCTCATATT 120
121 ACTTGGGCGCCCAATTTGGGTTTCAGTGTCTGTTTCCAGCTATGGGAGAGGTAAGTTA 180
121 ACTTGGGCGCCCAATTTGGGTTTCAGTGTCTGTTTCCAGCTATGGGAGAGGTAAGTTA 180
181 AACCAACCGTAATCG 197
181 AACCAACCGTAATCG 197

RESULT 4
AAC62380
ID AAC62380 standard; DNA; 198 BP.
XX
AC AAC62380;
XX
XX
DT 19-MAR-2001 (first entry)
XX
DE Leader region of a satellite tobacco mosaic virus (STMV).

XX Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
KW Gossypium; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
KW tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus; ss.
XX
OS Satellite tobacco mosaic virus.
XX
FN WO200063397-A2.
XX
XX
PD 26-OCT-2000.
XX
FF 17-APR-2000; 2000MO-EP003521.
XX

PR 20-APR-1999; 99US-00294022.
XX
PA (AVET) AVENTIS CROPS SCIENCE NV.
XX
PI Meulewaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaiff M;
XX
DR WPI; 2000-687182/67.
XX
XX Identifying and isolating genes involved in determining the trait or
PT phenotype of plant species, by infecting plants with gene silencing
PT constructs targeted to the gene, and identifying plants with altered
PT traits.
XX
PS Example 1; Page 63; 64pp; English.

The specification describes a method for isolating genes that determine a trait or phenotype of a plant species. The method comprises identifying a set of nucleic acids of genes correlated with the trait, creating a library of gene silencing constructs in a viral RNA vector, targeting a gene silencing constructs to the nucleic acid set, infecting a collection of individual plants with these, identifying plants with altered traits or phenotype, and isolating genes of the invention. The method is useful for isolating genes involved in the determination of trait or a phenotype of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypium, Triticum, Arabidopsis or Petunia. The method is also useful for modulating the expression of selected nucleic acid sequences and for validating the function of a nucleic acid sequence whose expression is correlated with the presence or absence of a specific trait in plants, but with otherwise unknown function. The method is also useful for developing agronomically useful products such herbicides or transgenic plants. The present sequence is a leader sequence of a satellite tobacco mosaic virus (STMV). The sequence is used to construct infective hybrid tobacco mosaic virus (TMV)/tobacco necrosis virus (TMV) vectors, for use in the method of the invention

Sequence 198 BP; 57 A; 41 C; 47 G; 53 T; 0 U; 0 Other;

Query Match 98.5%; Score 194; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.2e-57;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 AAATTTACCAATCAAAAGACCTTAACCAAGAGCTGTGCTGCTCATTTATGCTGTGGG 63
5 AAATTTACCAATCAAAAGACCTTAACCAAGAGCTGTGCTGCTCATTTATGCTGTGGG 64
64 GACATAGGGGAAAACATATGCTCTTCTTACAAGAGGCTTTCAGTGCCTCATATTACT 123
65 GACATAGGGGAAAACATATGCTCTTCTTACAAGAGGCTTTCAGTGCCTCATATTACT 124
124 TGGCGCCCAATTTGGGTTTCAGTGTCTGTTTCCAGCTATGGGAGAGGTAAGGTTAAAC 183
125 TGGCGCCCAATTTGGGTTTCAGTGTCTGTTTCCAGCTATGGGAGAGGTAAGGTTAAAC 184
184 CAAACCGTAATCG 197
185 CAAACCGTAATCG 198

RESULT 5
AAQ06316
ID AAQ06316 standard; DNA; 1059 BP.
XX
AC AAQ06316;
XX
XX 27-AUG-2003 (revised)
DT 06-FEB-1991 (first entry)
XX
XX Satellite Tobacco Mosaic Virus genomic sense RNA.
XX
XX Recombinant expression system; plant cell transformation;
KW Satellite Tobacco Mosaic Virus; STMV; ss.
XX
OS Satellite tobacco mosaic virus.

```
XX Key Location/Qualifiers
FH CDS 53..229
FT /*tag= a
FT /label= ORF 1
FT CDS 163..64
FT /*tag= b
FT /label= ORF 2
XX
XX W09012107-A.
XX
XX 18-OCT-1990.
XX
XX 31-MAR-1989; 89US-00332632.
XX
XX 31-MAR-1989; 89US-00332632.
XX
XX (SALK ) SALK INST BIOTECHN.
XX
XX Fitzmauric LC, Mirkov TE;
XX
XX WPI; 1990-334857/44.
XX
XX P-PSDB; AAR07475, AAR08254.
XX
XX Recombinant expression system used to transform plant cells - based on
XX satellite tobacco mosaic virus.
XX
XX Disclosure; Fig 1; 48pp; English.
XX
XX The two ORF have the potential code for proteins of 6.8Mr and 17.5Mr. The
XX 418 nucleotides following the ORF for the capsid do not encode any
XX polypeptides longer than 16 amino acids. The negative sense RNA has a
XX single ORF which has the coding capacity for an 8K protein. The sequence
XX as earlier published (B. Mirkov's Dissertation, cDNA Cloning, Nucleotide
XX Sequences, in vitro Translation, and Genome Organisation of STMV,
XX University of California, Riverside (March 1988)) contains several
XX incorrect nucleotides at and near the 3' and 5' termini, and erroneously
XX included extra nucleotides that resulted in a longer RNA molecule (1065
XX nt) than the viral RNA actually had (1059 nt). (Updated on 27-AUG-2003 to
XX correct OS field.)
XX
XX Sequence 1059 BP; 276 A; 226 C; 258 G; 0 T; 299 U; 0 Other;
XX
XX Query Match 98.5%; Score 194; DB 2; Length 1059;
XX Best Local Similarity 73.2%; Pred. No. 4.1e-57;
XX Matches 142; Conservative 52; Mismatches 0; Indels 0; Gaps 0;
XX
XX 4 AAACCTTACCAATCAAAAGACCTAACCAACAGGACTGTGCTGCTCATTTATGCTGTGGG 63
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 5 AAACUACCAUCAAAGACCUACCAACAGGACUGUGUGUUAUUGCUUGUGUGG 64
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 64 GACATAGGGGAAAACATATTGCTTCTTCTAAGAGGCGCTTCAGTCGCCATAATTA 123
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 65 GACAUAAGGGGAAAACAUUAUUGCUUCUUAAGAGGCGCUUCAGUGGCCAUUAUACU 124
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 124 TGGGCGCCAAATTTGGGTTTCAGTGTGCTTTCCAGCTATGCGGAGAGGTTAAAC 183
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 125 UGGCGCCCAAUUUGUGGUUUAUGUGUUGUUCUUCAGCUAUGGGGAGAGGUAAGG 184
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 184 CAAACCGTAATCG 197
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 185 CAAACCGUAUAUCG 198
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX
XX RESULT 6
XX ACC84999
XX ID ACC84999 standard; cDNA; 1059 BP.
XX
XX AC ACC84999;
XX
XX 13-OCT-2003 (first entry)
XX
XX STMV genome nucleotide sequence comprised in vector pSTMV-10.
```

```
XX Inhibitory RNA; viral RNA vector; coat protein; STMV; gene; ss.
XX Tobacco mosaic virus.
XX W02003052108-A2.
XX 26-JUN-2003.
XX
XX 05-DEC-2002; 2002WO-EP013964.
XX
XX 18-DEC-2001; 2001US-0340488P.
XX
XX (FARB ) BAYER BIOSCIENCE NV.
XX
XX Metzlauff MH, Gossele VML, Meulewaeter F, Fache ICA;
XX
XX WPI; 2003-523529/49.
XX
XX Introducing inhibitory RNA into a plant cell comprises providing a viral
XX RNA vector derived from a satellite RNA virus that encodes a coat
XX protein, and infecting a plant with the viral RNA vector and a
XX corresponding helper virus.
XX
XX Claim 4; Page 74-75; 86pp; English.
XX
XX The invention relates to introducing inhibitory RNA into a plant cell.
XX The method involves providing a viral RNA vector derived from a satellite
XX RNA virus having a sequence that encodes a coat protein, and infecting a
XX plant with the viral RNA vector and a corresponding helper virus. The
XX methods and viral RNA vectors are useful in introducing inhibitory RNA
XX into plant cells. These may be used to determine or validate the function
XX of isolated nucleic acid sequences in plants. The present sequence
XX represents the nucleotide sequence of the genome of STMV (satellite
XX tobacco mosaic virus), comprised in vector pSTMV-10. This sequence shows
XX a 5 nucleotide difference to that of the sequence (ACC85000) provided in
XX the GenBank
XX
XX Sequence 1059 BP; 276 A; 226 C; 258 G; 299 T; 0 U; 0 Other;
XX
XX Query Match 93.5%; Score 184.2; DB 9; Length 1059;
XX Best Local Similarity 95.9%; Pred. No. 1.1e-53;
XX Matches 189; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
XX
XX 1 AGTAAATCTTACCAATCAAAAGACCTAACCAACAGGACTGTGCTGCTCATTTATGCTGTTG 60
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 1 AGTAAATCTTACCAATCAAAAGACCTAACCAACAGGACTGTGCTGCTCATTTATGCTGTTG 60
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 61 GGGACATAGGGGAAAACATATTGCTTCTTCTAAGAGGCGCTTCAGTCGCCATAATT 120
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 61 GGGACATAGGGGAAAACATATTGCTTCTTCTAAGAGGCGCTTCAGTCGCCATAATT 120
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 121 ACTTGGCGCCCAATTTTGGGTTTCAGTGTGCTTTCCAGCTATGCGGAGAGGTAAGGTTA 180
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 121 ACTTGGCGCCCAATTTTGGGTTTCAGTGTGCTTTCCAGCTATGCGGAGAGGTAAGGTTA 180
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 181 AACCAACCGTAATCG 197
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX 181 AACCAACCGTAATCG 197
XX |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
XX
XX RESULT 7
XX AAQ06899/c
XX ID AAQ06899 standard; DNA; 62 BP.
XX
XX AC AAQ06899;
XX
XX 27-AUG-2003 (revised)
XX 06-FEB-1991 (first entry)
XX
XX Probe to detect STMV genomic sense RNA transformed E.coli.
XX Recombinant expression system; plant cell transformation; probe;
```


KW Satellite Tobacco Mosaic Virus; STMV; ss.
 OS Satellite tobacco mosaic virus.
 XX
 PN WO9012107-A.
 XX
 PD 18-OCT-1990.
 XX
 PF 31-MAR-1989; 89US-003332632.
 XX
 PR 31-MAR-1989; 89US-003332632.
 XX
 PA (SALK) SALK INST BIOTECHN.
 XX
 PI Fitzmauric LC, Mirkov TE;
 XX
 DR WPI; 1990-334857/44.
 DR
 DR P-PSDB; AAR07475, AAR08254.
 XX
 XX Recombinant expression system used to transform plant cells - based on
 PT satellite tobacco mosaic virus.
 PT
 XX
 PS Disclosure; Fig 1; 48pp; English.
 XX
 CC The probe is complementary to the 5' end of genomic STMV RNA. Clones
 CC identified with this probe were anticipated to contain full-length copies
 CC of the STMV genome. See also AAQ06316 for the STMV RNA sequence. (Updated
 CC on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 62 BP; 13 A; 10 C; 16 G; 23 T; 0 U; 0 Other;
 Query Match 24.8%; Score 48.8; DB 2; Length 62;
 Best Local Similarity 96.2%; Pred. No. 7.9e-07;
 Matches 50; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 AAACCTTACCATTCAAAAGACCTTAACCAACAGGACTGTGCTGCTATTATGC 55
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 52 AAACCTTACCATTCAAAAGACCTTAACCAACAGGACTGTGCTGCTATTATGC 1
 RESULT 8
 AAA29039/c
 ID AAA29039 standard; cDNA; 845 BP.
 XX
 AC AAA29039;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE B. napus flowering locus F (FLF) homologous cDNA 8.
 XX
 KW Flowering locus F; transcription factor; DNA binding domain; MADS box;
 KW FLF; chromosome 5; RFLP marker 447; flowering time; vernalization; ss.
 XX
 OS Brassica napus.
 XX
 FH Key Location/Qualifiers
 CDS 67..661
 FT /*tag= a
 FT /product= "FLF_homologue"
 XX
 PN WO200032780-A1.
 XX
 PD 08-JUN-2000.
 XX
 PF 02-DEC-1999; 99WO-AU001079.
 XX
 PR 03-DEC-1999; 98AU-00007469.
 PR
 PR 22-JAN-1999; 99US-0116928P.
 XX
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA (PERE/) PEREZ P.
 XX
 PI Perez P, Burn JF, Peacock WJ, Dennis ES, Sheldon CC;

PI Helliwell CA, Rouse DT;
 XX
 DR WPI; 2000-451762/39.
 DR
 XX P-PSDB; AAY96415.
 XX
 PT New nucleic acid comprising a MADS box that encodes a transcription
 PT factor for altering the flowering time or modifying the vegetative or
 PT floral phenotype of a plant such as wheat.
 XX
 PS Claim 6; Page 89-90; 106pp; English.
 XX
 CC AAA29032-41 are cDNA clones isolated from a Brassica napus cDNA library
 CC using hybridisation to an FLF (flowering locus F) cDNA from Arabidopsis
 CC thaliana. From both amino acid and nucleotide sequence data it was
 CC concluded that the clones represent transcripts from 5 different genes.
 CC The FLF gene encodes a transcription factor protein comprising a
 CC distinctive conserved DNA binding domain (a MADS box). The DNA was
 CC identified from a late-flowering mutant (flf), which arose during
 CC transformation of the early-flowering ecotype C24 with a modified binary
 CC vector pbin-delta-Ac. It was found that two T-DNAs segregate in the very
 CC late flowering phenotype and that two expressed genes (A and B) were
 CC closely linked to these T-DNAs. Gene B was re-designated as FLF. The FLF
 CC gene is linked to the top of chromosome 5, 4 cm from RFLP marker 447. The
 CC protein is capable of altering the flowering time of a plant. Over-
 CC expression of the FLF gene can be used to delay flowering in a plant.
 CC Alternatively, reduction of FLF gene expression, e.g. by antisense
 CC constructs, can be used to induce early flowering. Methods of modifying
 CC vegetative and/or floral phenotypes of a plant comprising increasing the
 CC level of expression of a FLF gene to modify the production or activity of
 CC a gibberellin in the plant are also claimed. The response of a plant to
 CC vernalization can also be altered. Any dicotyledonous or monocotyledonous
 CC plant species including horticultural plants, particularly those in the
 CC cut-flower industry, vegetable, cereal and other crop plants are capable
 CC of being controlled
 XX
 SQ Sequence 845 BP; 261 A; 183 C; 194 G; 207 T; 0 U; 0 Other;
 Query Match 17.4%; Score 34.2; DB 3; Length 845;
 Best Local Similarity 67.6%; Pred. No. 0.27;
 Matches 48; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 88 TTCTTCTACAGAGGCTTCAGTCGCCATTAATTAATTCCTGGCCCAATTTGGTTTCAGT 147
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 81 TTCTTCTCTCCCGCTTCTGCTCTCCACAGTCTTTGCGCCCTAATTTGATCTTGAGG 22
 QY 148 TGCTGTTTCCA 158
 DB ||||||||
 21 TCCGGTTTCCA 11
 RESULT 9
 AAA29035/c
 ID AAA29035 standard; cDNA; 868 BP.
 XX
 AC AAA29035;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE B. napus flowering locus F (FLF) homologous cDNA 4.
 XX
 KW Flowering locus F; transcription factor; DNA binding domain; MADS box;
 KW FLF; chromosome 5; RFLP marker 447; flowering time; vernalization; ss.
 XX
 OS Brassica napus.
 XX
 FH Key Location/Qualifiers
 CDS 80..673
 FT /*tag= a
 FT /product= "FLF_homologue"
 XX
 PN WO200032780-A1.
 XX
 PD 08-JUN-2000.

```
XX PF 02-DEC-1999; 99WO-AU001079.
XX PR 03-DEC-1998; 98AU-00007469.
XX PR 22-JAN-1999; 99US-0116928P.
XX PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX PA (PERE/) PEREZ P.
XX PI Perez P, Burn JE, Peacock WJ, Dennis ES, Sheldon CC;
XX PI Helliwell CA, Rouse DT;
XX DR WPI; 2000-451762/39.
XX DR P-PSDB; AAY96411.
XX PT New nucleic acid comprising a MADS box that encodes a transcription
XX PT factor for altering the flowering time or modifying the vegetative or
XX PT floral phenotype of a plant such as wheat.
XX PS Claim 6; Page 87-88; 106pp; English.
XX CC AA29032-41 are cDNA clones isolated from a Brassica napus cDNA library
XX CC using hybridisation to an FLF (flowering locus F) cDNA from Arabidopsis
XX CC thaliana. From both amino acid and nucleotide sequence data it was
XX CC concluded that the clones represent transcripts from 5 different genes.
XX CC The FLF gene encodes a transcription factor protein comprising a
XX CC distinctive conserved DNA binding domain (a MADS box). The DNA was
XX CC identified from a late-flowering mutant (flf), which arose during
XX CC transformation of the early-flowering ecotype C24 with a modified binary
XX CC vector pBin-delta-Ac. It was found that two T-DNAs segregate in the very
XX CC late flowering phenotype and that two expressed genes (A and B) were
XX CC closely linked to these T-DNAs. Gene B was re-designated as FLF. The FLF
XX CC gene is linked to the top of chromosome 5, 4 cm from RFLP marker 447. The
XX CC protein is capable of altering the flowering time of a plant. Over-
XX CC expression of the FLF gene can be used to delay flowering in a plant.
XX CC Alternatively, reduction of FLF gene expression, e.g. by antisense
XX CC constructs, can be used to induce early flowering. Methods of modifying
XX CC vegetative and/or floral phenotypes of a plant comprising increasing the
XX CC level of expression of a FLF gene to modify the production or activity of
XX CC a gibberellin in the plant are also claimed. The response of a plant to
XX CC vernalization can also be altered. Any dicotyledonous or monocotyledonous
XX CC plant species including horticultural plants, particularly those in the
XX CC cut-flower industry, vegetable, cereal and other crop plants are capable
XX CC of being controlled
XX SQ Sequence 868 BP; 273 A; 187 C; 196 G; 212 T; 0 U; 0 Other;
Query Match 17.4%; Score 34.2; DB 3; Length 868;
Best Local Similarity 67.6%; Pred. No. 0.27;
Matches 48; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 88 TTCTTCTACAGAGCCCTTCAGTCGCCATAATTAATCTTGGCGCCCAATTTGGGTTTCAGT 147
Db 93 TTCTTCTCTCCCATGGCTTCTGTCTCCACACAGTGTCTTGGCGCCCTAATTTGATCTTGAGG 34
Qy 148 TGCTGTTTCCA 158
Db 33 TCCGGTTTCCA 23
RESULT 10
AAA29041/C
ID AAA29041 standard; cDNA; 891 BP.
XX AC AAA29041;
XX DT 12-SEP-2000 (first entry)
XX DE B. napus flowering locus F (FLF) homologous cDNA 10.
XX KW Flowering locus F; transcription factor; DNA binding domain; MADS box;
XX KW FLF; chromosome 5; RFLP marker 447; flowering time; vernalization; ss.
XX
```

```
OS Brassica napus.
XX Key Location/Qualifiers
XX CDS 67..660
XX FT /*tag= a
XX FT /product= "FLF_homologue"
XX PN WO200032780-A1.
XX PD 08-JUN-2000.
XX XX 02-DEC-1999; 99WO-AU001079.
XX XX 03-DEC-1998; 98AU-00007469.
XX PR 22-JAN-1999; 99US-0116928P.
XX XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX PA (PERE/) PEREZ P.
XX PI Perez P, Burn JE, Peacock WJ, Dennis ES, Sheldon CC;
XX PI Helliwell CA, Rouse DT;
XX DR WPI; 2000-451762/39.
XX DR P-PSDB; AAY96417.
XX PT New nucleic acid comprising a MADS box that encodes a transcription
XX PT factor for altering the flowering time or modifying the vegetative or
XX PT floral phenotype of a plant such as wheat.
XX PS Claim 6; Page 90-91; 106pp; English.
XX CC AA29032-41 are cDNA clones isolated from a Brassica napus cDNA library
XX CC using hybridisation to an FLF (flowering locus F) cDNA from Arabidopsis
XX CC thaliana. From both amino acid and nucleotide sequence data it was
XX CC concluded that the clones represent transcripts from 5 different genes.
XX CC The FLF gene encodes a transcription factor protein comprising a
XX CC distinctive conserved DNA binding domain (a MADS box). The DNA was
XX CC identified from a late-flowering mutant (flf), which arose during
XX CC transformation of the early-flowering ecotype C24 with a modified binary
XX CC vector pBin-delta-Ac. It was found that two T-DNAs segregate in the very
XX CC late flowering phenotype and that two expressed genes (A and B) were
XX CC closely linked to these T-DNAs. Gene B was re-designated as FLF. The FLF
XX CC gene is linked to the top of chromosome 5, 4 cm from RFLP marker 447. The
XX CC protein is capable of altering the flowering time of a plant. Over-
XX CC expression of the FLF gene can be used to delay flowering in a plant.
XX CC Alternatively, reduction of FLF gene expression, e.g. by antisense
XX CC constructs, can be used to induce early flowering. Methods of modifying
XX CC vegetative and/or floral phenotypes of a plant comprising increasing the
XX CC level of expression of a FLF gene to modify the production or activity of
XX CC a gibberellin in the plant are also claimed. The response of a plant to
XX CC vernalization can also be altered. Any dicotyledonous or monocotyledonous
XX CC plant species including horticultural plants, particularly those in the
XX CC cut-flower industry, vegetable, cereal and other crop plants are capable
XX CC of being controlled
XX SQ Sequence 891 BP; 276 A; 191 C; 202 G; 222 T; 0 U; 0 Other;
Query Match 17.4%; Score 34.2; DB 3; Length 891;
Best Local Similarity 67.6%; Pred. No. 0.27;
Matches 48; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 88 TTCTTCTACAGAGCCCTTCAGTCGCCATAATTAATCTTGGCGCCCAATTTGGGTTTCAGT 147
Db 80 TTCTTCTCTCCCATGGCTTCTGTCTCCACACAGTGTCTTGGCGCCCTAATTTGATCTTGAGG 21
Qy 148 TGCTGTTTCCA 158
Db 20 TCCGGTTTCCA 10
RESULT 11
ABK84699/c
ID ABK84699 standard; cDNA; 198285 BP.
```

XX AC ABK84699;
XX DT 14-AUG-2002 (first entry)
XX DE Human cDNA differentially expressed in granulocytic cells #1270.
XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX OS Homo sapiens.
XX PN WO200228999-A2.
XX PD 11-APR-2002.
XX PF 03-OCT-2001; 2001WO-US030821.
XX PR 03-OCT-2000; 2000US-0237189P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX DR Detecting granulocyte activation by detecting differential expression of
XX PT genes associated with granulocyte activation, which serves as diagnostic
XX PT markers that is useful for monitoring disease states and drug toxicity.
XX PS Claim 1; SEQ ID NO 1270; 114pp; English.
XX CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;
XX Query Match 16.9%; Score 33.2; DB 6; Length 198285;
XX Best Local Similarity 51.3%; Pred. No. 4.7;
XX Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 35 GACTGTCTGGTCAATTTATGCTGTGGGGGACATAGGGGAAACATATTCCTTCTCT 94
DB 162395 GATTCTCTGGGTATTTGAAGATCCTTGGGGACAAATGCCAAGAAGACACCAACTTCTTT 162336
QY 95 ACRAAGAGCCCTTCAGTCGCCATAATTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTT 154
DB 162335 ATCAGCATCTCTCAGAAATCCATGCCCTGTGGGTAGTTATATGTGATATCAGAAATCGCCC 162276
QY 155 TCCAGCTATGGGAGAGGTAAGGTTAAACC 184
DB 162275 TCCACTCCAGGAGACCCCAAGGCTTAATC 162246
RESULT 12
ABN97319/c
ID ABN97319 standard; DNA; 198285 BP.
XX AC ABN97319;
XX DT 13-AUG-2002 (first entry)
XX DE Gene #3817 used to diagnose liver cancer.
XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumor; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US030589.
XX PR 02-OCT-2000; 2000US-0237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX DR Diagnosing and detecting the progression of liver cancer, hepatocellular
XX PT carcinoma or metastatic liver tumor in a patient, involves detecting the
XX PT level of expression of two or more genes in a liver tissue sample.
XX PS Claim 1; SEQ ID NO 3817; 298pp; English.
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumor in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cytostatic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
XX CC for this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;

Query Match 16.9%; Score 33.2; DB 6; Length 198285;
Best Local Similarity 51.3%; Pred. No. 4.7; Indels 0; Gaps 0;
Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
SQ Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;

QY 35 GACTGTCGTGTCATTTATGCTGTGGGGGACATAGGGGAAACATATTGCTTCTTCT 94
Db 162395 GATTCTCTGGGTATTGGAAGATCCTTGGGACAAATGCCAAGAGACACCAACTTCTTTT 162336

QY 95 ACAAGAGCCCTTCAGTCGCCATAATTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTT 154
Db 162335 ATCAGCATCTCTCAGATCCATGCCCTGTGGGTAGTTATGTGATATCAGATCGCC 162276

QY 155 TCCAGCTATGGGAGAGGTAAAGGTTAAACC 184
Db 162275 TCCACTGCCAGGAGACCCAAAGGCTTAATC 162246

RESULT 13
ADRS52987/c
ID ADRS52987 standard; DNA; 198285 BP.
XX AC ADRS2987;
XX 18-NOV-2004 (first entry)
XX Drug therapy altered expressed gene #338.
XX drug activity monitoring; expression profile; gene expression;
KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW mTOR; ds.
XX Homo sapiens.
XX WO2004072265-A2.
XX 26-AUG-2004.
XX 11-FEB-2004; 2004WO-US004118.
XX 11-FEB-2003; 2003US-0446133P.
XX 03-APR-2003; 2003US-0459782P.
XX 23-JAN-2004; 2004US-0538246P.
XX (AMHP) WYETH.
XX (BURC) BURCZYNSKI M.
XX (TWIN) TWINE N.
XX (DORN) DORNER A. J.
XX (TREP) TREPICCHIO W. L.
XX Burczynski M, Twine N, Dornier AJ, Trepicchio WL;
FI WPI; 2004-642301/62.
XX
XX Monitoring drug activities in vivo comprises comparing an expression
PT profile of a gene in a peripheral blood sample of a patient before and
PT after drug therapy.
XX
XX Disclosure; SEQ ID NO 338; 136pp; English.
XX
XX The invention relates to a method of monitoring drug activities in vivo
CC by comparing an expression profile of at least one gene in a peripheral
CC blood sample of a patient to a reference expression profile of the at
CC least one gene, where the at least one gene is differentially expressed
CC in peripheral blood mononuclear cells (PBMCs) of patients who have a non-
CC blood disease and are subjected to a drug therapy as compared to PBMCs
CC isolated from the patient before the drug therapy, and where the patient
CC has the non-blood disease and is being treated by the drug therapy. The
CC method, kit, and nucleic acid array are useful for monitoring drug
CC activities in vivo. The drug is especially CCI-779, an ester analogue of
CC the immunosuppressant rapamycin which is a potent inhibitor of the
CC mammalian target of rapamycin (mTOR). This sequence represents a gene

CC expressed in PBMC altered by the drug therapy. (Note: this sequence does
CC no form part of the printed specification but was obtained in electronic
CC format from WIPO at ftp.wipo.int/pub/published_pct_sequences/).

QY 35 GACTGTCGTGTCATTTATGCTGTGGGGGACATAGGGGAAACATATTGCTTCTTCT 94
Db 162395 GATTCTCTGGGTATTGGAAGATCCTTGGGACAAATGCCAAGAGACACCAACTTCTTTT 162336

QY 95 ACAAGAGCCCTTCAGTCGCCATAATTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTT 154
Db 162335 ATCAGCATCTCTCAGATCCATGCCCTGTGGGTAGTTATGTGATATCAGATCGCC 162276

QY 155 TCCAGCTATGGGAGAGGTAAAGGTTAAACC 184
Db 162275 TCCACTGCCAGGAGACCCAAAGGCTTAATC 162246

RESULT 14
ACN44898/c
ID ACN44898 standard; DNA; 144723 BP.
XX AC ACN44898;
XX 18-NOV-2004 (first entry)
XX Human genomic sequence hCG18129.
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX Homo sapiens.
XX WO2003073826-A2.
XX 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006235.
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 1576; Opp; English.
XX
XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US200218586A1, for which no sequence data was published
XX
XX Sequence 144723 BP; 41373 A; 28002 C; 30266 G; 45082 T; 0 U; 0 Other;
SQ

Query Match 15.5%; Score 30.6; DB 11; Length 144723;
Best Local Similarity 52.8%; Pred. No. 34;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 72 GGAACAATATGCTTCTCTACAAGAGGCTTTCAGTCGCCATAATTACTTGGCGCC 131
DB 24352 GAGTACTCATTCAGTCTCTCTGTAAAGAGCTTTAAAAAGCTCATTTTGTGCTTT 24293

QY 132 AATTTGGGTTTCAGTCTCTGTTTCCAGCTATGGGAGAGGTAAGGTTAAACCAACCGT 191
DB 24292 GTTTTGGTCAGAGATCTTTTCCCTTACTGAAACTGAGAGTATTAAACCAACAC 24233

QY 192 AAATC 196
DB 24232 AAATC 24228

RESULT 15
ACN84908/C
ID ACN84908 standard; DNA; 819 BP.

XX AC ACN84908;
XX
XX
XX 02-DEC-2004 (first entry)
XX
XX Breast cancer related marker, seq id 6058.
XX
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX Homo sapiens.
XX
XX US2003099974-A1.
XX
XX 29-MAY-2003.
XX
XX 18-JUL-2002; 2002US-00198846.
XX
XX 18-JUL-2001; 2001US-0306220P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2003-787014/74.
XX
XX Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
XX Disclosure; SEQ ID NO 6058; 36pp; English.

The invention relates to an isolated polypeptide (I) associated with
breast cancer which is encoded by a nucleic acid molecule comprising a
nucleotide sequence (S1). Further disclosed is an antibody that binds to
the polypeptide of the invention. The activity of the polypeptide of the
invention may be described as cytostatic. The antibody is useful for
detecting the presence of (I) in a sample. Nucleic acid molecules of the
invention are useful in the detection of breast tumours. (I) is useful as
a marker for breast cancer and in breast cancer therapy. Sequences given
in records ACN78851-ACN92934 represent nucleic acid markers associated
with breast cancer. Note: The sequence listing does not form part of the
specification but may be obtained in electronic format from the USPTO web
site at seqdata.uspto.gov/sequence.html?docID=20030099974

Query Match 15.4%; Score 30.4; DB 11; Length 819;
Best Local Similarity 50.5%; Pred. No. 5.6;
Matches 55; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 73 GGAACAATATGCTTCTCTACAAGAGGCTTTCAGTCGCCATAATTACTTGGCGCCA 132
DB 413 GGACCAANTATTGCGATGTTNNNNAGAGCNCATTTTCCCTCATTTAAAGGGNACANA 354

QY 133 ATTTTGGGTTTTCAGTTCGTGTTTCCAGCTATGGGAGAGGTAAGGTTAA 181
DB 353 ATCATGGGTACCGGGCCCCCCTAGAGGTAGCGGTATCGATAAGCTTAA 305

Search completed: September 26, 2005, 01:44:57
Job time : 252.69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 00:52:04 ; Search time 1595.3 Seconds
(without alignments)
4700.480 Million cell updates/sec

Title: US-09-551-494-4_COPY_1_197

Perfect score: 197

Sequence: 1 agtaaaacttaacatacaaaa.....ttaaaccaaacgcgtataatcg 197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	35.4	18.0	1056	8	CC253821 CH261-136
2	34.2	17.4	922	9	CU145583 ISB1-146A
3	34	17.3	623	2	BB981434 UI-M-CG0p
4	33.4	17.0	664	9	CL386218 RPCI44_33
5	33.4	17.0	955	9	AL097736 Drosophil
6	33	16.8	597	8	AZ214225 Sheared D
7	33	16.8	618	9	AL015499 F.rubripe
8	32.6	16.5	677	1	AL730822 AL730822
9	32.6	16.5	741	1	AL715244 AL715244
10	32.6	16.5	754	9	AG491548 Mus muscu
11	32.6	16.5	841	7	CV470085 42879.1 C
12	32.4	16.4	428	8	AQ600889 HS_5317 A
13	32.4	16.4	473	8	AQ240675 CIT-HSP-2
14	32.4	16.4	945	7	CK405143 AUF Ifspn
15	32.2	16.3	170	1	AI2335685 EST232247
16	32.2	16.3	452	8	AQ611644 HS_5085 B
17	32.2	16.3	596	1	AI883305 fc60b09.Y
18	32	16.2	1209	9	BI683952 603306789
19	32	16.2	1242	9	AG323879 Mus muscu
20	31.8	16.1	518	8	BZ833894 CH240_236
21	31.8	16.1	925	7	CK182690 EST772005
22	31.6	16.0	389	7	T29967 EST100848 H
23	31.6	16.0	453	4	BW433343 lOm12H11
C 24	31.6	16.0	1196	8	CC185235 CH261-30F

C 25	31.6	16.0	2205	9	CL982354	CL982354	ObIFSC047
C 26	31.4	15.9	320	1	AI317531	AI317531	uj23h01.Y
C 27	31.4	15.9	532	8	AZ153267	AZ153267	SP_0051.B
C 28	31.4	15.9	712	9	AG286153	AG286153	Mus_muscu
C 29	31.2	15.8	704	5	BF008523	BF008523	BP008523
C 30	31.2	15.8	705	6	CB941634	CB941634	AGENCOURT
C 31	31.2	15.8	705	9	CL354887	CL354887	RPCI44_40
C 32	31	15.7	398	6	CD825381	CD825381	BN25_060J
C 33	31	15.7	444	8	AZ993747	AZ993747	2M0278J22
C 34	31	15.7	649	2	AW697349	AW697349	NF115F09S
C 35	31	15.7	679	4	BI959305	BI959305	HVSMEN001
C 36	31	15.7	753	1	AJ817392	AJ817392	AKh07f11.
C 37	31	15.7	756	8	BZ069163	BZ069163	1kh07f11.
C 38	31	15.7	766	9	CR837330	CR837330	GR0AAA71D
C 39	31	15.7	797	6	CA466704	CA466704	AGENCOURT
C 40	30.8	15.6	322	9	CG670009	CG670009	OST467883
C 41	30.8	15.6	366	6	CA521970	CA521970	KS12001B1
C 42	30.8	15.6	601	6	CD903073	CD903073	G356_109A
C 43	30.8	15.6	732	7	CO541510	CO541510	gm_A002.6
C 44	30.8	15.6	885	5	BO441171	BO441171	AGENCOURT
C 45	30.8	15.6	900	8	BH166924	BH166924	ENTRV64TF

ALIGNMENTS

RESULT 1
CC253821/c
LOCUS
DEFINITION CC253821 1056 bp DNA linear GSS 13-MAY-2003
CH261-136J24 Sp6.2 CH261 Gallus gallus genomic clone CH261-136J24,
genomic survey sequence.
ACCESSION CC253821
VERSION CC253821.1 GI:30590571
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 17
High quality sequence stop: 657.
FEATURES
source
1..1056
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-136J24"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 18.0%; Score 35.4; DB 8; Length 1056;
Best Local Similarity 55.2%; Pred No. 3.2;
Matches 69; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
Qy 69 AGGGGAAACATATTCCTTCTTACAGAGGCGCTTACGCCCAATTAATTCTTGGC 128

```

Db      1054 AGGAGAGAAAAATAGTCTGAGTTATTTATATAAAGAGTTGGAGAGAAACAAAAAGGCGTCAAGG 995
QY      129 CCCAATTTTGGGTTTCAGTTGCTGTTTCCAGCTATGGGAGAGTAAGGTTAAACCAAC 188
Db      994 TTTTATTTGGGGTTTGGCGGTGTTTATTTAGGGATAAGAGAAATGTTTTTTTTTAACCAAC 935
QY      189 CGTAA 193
Db      934 AGAAA 930

RESULT 2
CL145583
LOCUS      CL145583
DEFINITION ISB1-146A20_Sp6.1 ISB1 Xenopus tropicalis genomic clone
ACCESSION CL145583
VERSION    CL145583.1 GI:40639218
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 922)
AUTHORS    Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 14
            High quality sequence stop: 715.
FEATURES   source
            1..922
                Location/Qualifiers
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone_lib="ISB1-146A20"
                /clone_lib="ISB1"
                /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
Query Match      17.4%; Score 34.2; DB 9; Length 922;
Best Local Similarity 71.4%; Pred. No. 7.6;
Matches 45; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY      124 TGGCGCCCAATTTTGGGTTTCAGTTGCTGTTTCCAGCTATGGGAGAGGTAAGGTTAAAC 183
Db      792 TGGGGGCCAATTTTGGTTTAACTTTCGTTTAACTTAAGGGGAAGGTTATTTATAAC 851
QY      184 CAA 186
Db      852 CAA 854

RESULT 3
BE981434
LOCUS      BE981434
DEFINITION UI-M-CG0p-bdc-b-05-0-UI.s1 NIH BMAP Ret4_S2 Mus musculus cDNA clone
ACCESSION BE981434
VERSION    BE981434.1 GI:10650537
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 623)
AUTHORS    Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    97044477
PUBMED     889548
COMMENT    Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
            clones from RESEARCH GENETICS. It should be noted that Bento Soares
            is generating a small number of additional specialized
            non-redundant arrays of BMAP cDNAs whose availability will be
            considered under appropriate and limited collaborative arrangements
            The tissue for this library was contributed by Dr. Xin-Yuan Fu,
            Yale University School of Medicine
            Seq primer: M13 Forward
            POLYA=No.
FEATURES   Location/Qualifiers
            1..623
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UI-M-CG0p-bdc-b-05-0-UI"
                /lab_host="DHI0B (Life Technologies)"
                /clone_lib="NIH BMAP Ret4_S2"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker. Site 1: Not I; Site 2: Eco RI; The
                NIH BMAP Ret4_S2 library is a subtracted library,
                ultimately derived from mouse retina tissue libraries at
                various stages of development. For a detailed description
                of the library from which this clone was derived, please
                visit our web site at brainest.eng.uiowa.edu. The tissue
                for this library was contributed by Dr. Xin-Yuan Fu, Yale
                University School of Medicine
                TAG_SEQ=None found"
ORIGIN
Query Match      17.3%; Score 34; DB 2; Length 623;
Best Local Similarity 57.5%; Pred. No. 8.2;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY      89 TCTTCTACAAGAGCCCTTCAGTCGCCATAATTACTTGGCGCCCAATTTTGGGTTTCAGTT 148
Db      458 TCTACCAGAAGTCGCTTTAGAAAGACGCGAGTTACTTGGAGCAGAAATTTGTTGTCAAGAT 517
QY      149 GCTGTTTCCAGCTATGGGGAGAGTAAGGTTAAACCAACCGTAAA 194
Db      518 TATGCTTTTGACTTATATGAAGTTGAATTGACAAACAAACGAAGAGA 563

RESULT 4
CL386218
LOCUS      CL386218
DEFINITION RPCI44_333C6.r RPCI-44 Sus scrofa genomic clone RPCI44_333C6,
            genomic survey sequence.
ACCESSION CL386218
VERSION    CL386218.1 GI:51438183
KEYWORDS   GSS.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```


REFERENCE 1 (bases 1 to 664)
 AUTHORS Begatcheva,M.B., Meyers,S., He,W., Larkin,D.M., Marron,B.M.,
 Beaver,J.B. and Schook,L.B.
 TITLE Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
 JOURNAL Through Comparative Genomics
 COMMENT Unpublished (2004)
 Other GSSs: RPCI44_333C6.f
 Contact: Lawrence B. Schook
 Department of Animal Sciences
 University of Illinois at Urbana Champaign
 1201 W. Gregory Dr., Urbana, IL 61801, USA
 Tel: 217 265 5326
 Fax: 217 244 5617
 Email: schook@uiuc.edu
 Clones are derived from the porcine BAC library RPCI-44
 (http://www.bacpac.choxi.org/porcine242.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@choxi.org).
 Clones may be purchased from BACPAC Resources
 (http://BACPACorders.choxi.org). This work was undertaken as part
 of the International Swine Genome Sequencing Consortium by
 University of Illinois at Urbana Champaign, USA with funds provided
 by grant No. AG2002-34480-11828 from USDA-CSREES and
 AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing
 Initiative)
 Plate: 333 row: C column: 6
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 Location/Qualifiers
 1..664
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /strain="four pigs (breed: 37.5% Yorks Landrace and 25%
 Meishan)"
 /db_xref="taxon:9823"
 /clone="RPCI44_333C6"
 /sex="male"
 /cell_type="blood"
 /clone_lib="RPCI-44"
 /note="Vector: pTRABAC2; Site 1: EcoRI; Site 2: EcoRI;
 porcine male BAC library produced by Pieter de Jong"

ORIGIN
 Query Match 17.0%; Score 33.4; DB 9; Length 664;
 Best Local Similarity 57.0%; Pred. No. 13;
 Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 74 GAAACATATGCTCTCTTACAGAGGCTTCAGTCGCATATTAATCTTGGCGCCAA 133
 Db 505 GAAAGTTCTCTCTCTTACATGCTACTAGTATTCATTTGAGGTGACAC 564

Qy 134 TTTGGGTTTCAGTTGCTGTTTCCAGCTATGGGAGAGTAAGTTA 180
 Db 565 GTTTGGGTTTATTTGTTTCTTACTATGATTACAAGTAATGTA 611

RESULT 5
 CNS002J2 955 bp DNA linear GSS 28-JUL-1999
 LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
 DEFINITION BACN02F06 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCSSION AL097736
 VERSION AL097736.1 GI:5609347
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 955)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : segr@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (BDGP) -
 http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobAC11.

FEATURES
 source
 Location/Qualifiers
 1..955
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACN02F06"
 /clone_lib="DrosBAC"
 /plasmid="pBelobAC11"
 /note="end : SP6"

ORIGIN
 Query Match 17.0%; Score 33.4; DB 9; Length 955;
 Best Local Similarity 33.7%; Pred. No. 14;
 Matches 57; Conservative 39; Mismatches 73; Indels 0; Gaps 0;

Qy 8 TTACCAATCAAAGACCTAACCAACAGAGCTGCTGCTGCTATTTATGCTGTTGGGGACA 67
 Db 920 KTHCCAAKGTCTTTGGYGCMATATKWTKWSGMMATKGTCTTTTWTWATGBAACCCG 861

Qy 68 TAGGGGAAAACATATTCCTCTTCTACAAGAGGCTTCAGTCGCATATTAATTTGCTTGGC 127
 Db 860 RAGRAGAWGASCAAGMTDTTAWRTACTTTCWMTGTRGSCMCATCCCMAYTGSYWCYVGA 801

Qy 128 GCCCAATTTGGGTTTCAGTTGCTGTTTCCAGCTATGGGAGAGGTAAG 176
 Db 800 GGCWCKTGGMATDACAATTGTGTTGGGVCCTASGSRACMARATAG 752

RESULT 6
 AZ214225 597 bp DNA linear GSS 09-JUN-2000
 LOCUS Sheared DNA-108G3.TR Sheared DNA Trypanosoma brucei genomic clone
 DEFINITION Sheared DNA-108G3, genomic survey sequence.
 ACCSSION AZ214225
 VERSION AZ214225.1 GI:8432025
 KEYWORDS GSS.
 SOURCE Trypanosoma brucei
 ORGANISM Trypanosoma brucei
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.
 REFERENCE 1 (bases 1 to 597)
 AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
 Gerrard,C., Leach,V., de Jong,P., Ullu,E., Melville,S.,
 Donelson,J., Fraser,C. and Adams,M.
 TITLE Determination of clone end sequences from Trypanosoma brucei GUTat
 10.1 sheared DNA library
 JOURNAL Unpublished (1999)
 COMMENT Other GSSs: Sheared DNA-108G3.TF
 Contact: Najib M. El-Sayed
 Department of Eukaryotic Genomes
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: nelsayed@tigr.org
 Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
 DNA library constructed at TIGR. Clones will be available for
 distribution through Research Genetics, Alabama, USA. Sheared DNA
 end sequences search page: http://www.tigr.org/tcdb/mdb/tcdb/.
 Seq primer: M13-Reverse
 Class: shotgun.
 Location/Qualifiers
 source 1..597

```

/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927/4 GUTat 10.1"
/db_xref="taxon:5691"
/clone_lib="Sheared DNA-108G3"
/clone_lib="Sheared DNA"
/notes="Vector: pUC18; Site 1: SmaI; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + i method used for the library construction is
described in detail in Smith, H.O. and Venter, J.C.
(Making small insert libraries for whole genome shotgun
sequencing projects. In Genome Sequencing: A Practical
Approach, eds. M. Vaudin and B. Borell, Oxford University
Press, 1999)."

```

```

ORIGIN
Query Match          16.8%; Score 33; DB 8; Length 597;
Best Local Similarity 51.7%; Pred. No. 17;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 51 TATCTCTGTTGGGACATAGGGGAAACACATATTGCTCTCTTACAGAGGCGCTTCAGT 110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 TCTCTCTTGTCTGCCACAGTATCAATATGCTGCTCTTAAATATCTTTCGGA 297

Qy 111 CGCATATTAATCTTTGGGCGCCCAATTTGGGTTTCAGTTCGTTTCCAGCATGGGAGA 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 TAATATTATATATTGTTCTCGCCATTTTAGCGATTGTACAACTTCTGCTTTGAGGAAG 357

Qy 171 GGTAAAGGTTAAACCAACCGTAAT 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 GTTGACATCAATAATAACGTAACAT 382

```

```

RESULT 7
LOCUS      FR0022633          618 bp DNA linear GSS 25-FEB-2004
DEFINITION F.rubripes GSS sequence, clone 079F10aG9, genomic survey sequence.
ACCESSION  AL015499
VERSION     AL015499.1 GI:2681867
KEYWORDS   GSS: genome survey sequence.
SOURCE     Takifugu rubripes (Fugu rubripes)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE  1 Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
            Bouchireb,N., Cottage,A., Yeo,G.S., Umrantia,Y., Williams,G. and
            Brenner,S.
            Generation and analysis of 25 Mb of genomic DNA from the pufferfish
            Fugu rubripes by sequence scanning
            Genome Res. 9 (10), 960-971 (1999)
JOURNAL    9945097
MEDLINE    10523524
REFERENCE  2 (bases 1 to 618)
            Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrantia,Y.,
            Williams,G. and Brenner,S.
            Direct Submission
            Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
            Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT    Vector: pBluescript II KS
            V type: phagemid
            PRIMER: KS
DESCR:     One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
FEATURES   source      Location/Qualifiers
            1..618
            /organism="Takifugu rubripes"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="079F10aG9"
/clone_lib="Cosmid 079F10"

ORIGIN
Query Match          16.8%; Score 33; DB 9; Length 618;
Best Local Similarity 50.4%; Pred. No. 17;
Matches 69; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 2  GTAACATTACCAATCAAAAGACCTTAACCAACAGGACTGTGTGTGCTATTATGCTGTGG 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 71  GNAACNNNGAGCTCAANCTGTAATTAATCTCAGCACNCGAGGTAGATGTGTATGATGCTCG 130

Qy 62  GGACATAGGGGGAACACATATTGCTCTTCTTACAGAGGCGCTTCAGTCGCCATAATTA 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 TGGAAAGAGGGGCAAAATCAATTTCCAGTGAGGCAAGAGGCGCTTCCAAACCATCAATACAN 190

Qy 122 CTTGGGCGCCCAATTTTG 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 GTCGACACCCAAACCTG 207

RESULT 8
LOCUS      AL730822          677 bp mRNA linear EST 18-APR-2002
DEFINITION AL730822 Danio rerio embryonic inner ear subtracted cDNA Danio
            rerio cDNA clone EN0AA116ZB06 5', mRNA sequence.
ACCESSION  AL730822
VERSION     AL730822.1 GI:20195426
KEYWORDS   EST.
SOURCE     Danio rerio (zebrafish)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE  1 (bases 1 to 677)
            Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M.,
            Hardelin,J.P., Weissenbach,J. and Petit,C.
            A subtracted cDNA library from the zebrafish (Danio rerio)
            embryonic inner ear
            Unpublished (2002)
COMMENT    Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
            Location/Qualifiers
            1..677
            /organism="Danio rerio"
            /mol_type="mRNA"
            /db_xref="taxon:7955"
            /clone="EN0AA116ZB06"
            /tissue_type="inner ear"
            /dev stage="embryonic"
            /clone_lib="Danio rerio embryonic inner ear subtracted
            cDNA"
            /note="subtracted cDNA library"

```

```

ORIGIN
Query Match          16.5%; Score 32.6; DB 1; Length 677;
Best Local Similarity 53.5%; Pred. No. 24;
Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 59  TGGGGGACATAGGGGGGAAAACATATTGCTCTTCTTACAGAGGCGCTTCAGTCGCCATAA 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 42  TGTGCCACATGGGGGAAAAAATAAACGGAATTGAGTCACCTTGAACCATGTCAGTGTAAATAA 101

Qy 119 TTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTCTTCCAGCTATGGGAGAGGTAAGGT 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 AGCCTAAACGAGCGCTTGGAGGCTTGAGACGCTCTAAACAGGTTCTGGGAGGTCATATAC 161

Qy 179 TAAACCA 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      162 TGAACCA 168

RESULT 9
AL715244
LOCUS   AL715244 Danio rerio embryonic inner ear subtracted cDNA Danio
DEFINITION
AL715244 741 bp mRNA linear EST 18-APR-2002
COMMENT  rerio cDNA clone BNOAA012ZF04 3', mRNA sequence.
VERSION  AL715244
KEYWORDS EST.
SOURCE   Danio rerio (zebrafish)
ORGANISM Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 741)
AUTHORS  Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M.,
          Hardelin,J.P., Weissenbach,J. and Petit,C.
TITLE    A subtracted cDNA library from the zebrafish (Danio rerio)
          embryonic inner ear
JOURNAL  Unpublished (2002)
COMMENT  Contact: Genoscope
          Genoscope - Centre National de Sequencage
          2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
          Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source   1..741
          /organism="Danio rerio"
          /mol_type="mRNA"
          /db_xref="taxon:7955"
          /clone="BNOAA012ZF04"
          /tissue_type="inner ear"
          /dev_stages="embryonic"
          /clone_lib="Danio rerio embryonic inner ear subtracted
          cDNA"
          /note="subtracted cDNA library"

ORIGIN
Query Match      16.5%; Score 32.6; DB 1; Length 741;
Best Local Similarity 53.5%; Pred. No. 24;
Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy  59 TGGGGGACATAGGGGAAAACATATTGCTCTTCTCAAGAGGCTTCAGTCGCCATAA 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  42 TGTGCCACATGGGAAAAATAAACGGAATTGAGTCACCTTGAAACCATGCACTGTAATAA 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  119 TTACTTGGCGCCCAATTTTGGGTTTCAGTGTGTTTCCAGCTATGGGGAGGTAAGGT 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  102 AGCCTAAACGAGCGCTTGGAGGCTTGAGACGCTCTAAACAGTTCTGGGAGGTCAATATAC 161

Qy  179 TAAACCA 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  162 TGAACCA 168

RESULT 10
AG491548/c
LOCUS   AG491548 754 bp DNA linear GSS 04-JUN-2004
DEFINITION
AG491548 Mus musculus molossinus DNA, clone:MSMg01-389D22.TJ, genomic survey
sequence.
ACCESSION AG491548
VERSION   AG491548.1 GI:48198778
KEYWORDS GSS.
SOURCE   Mus musculus molossinus
ORGANISM Mus musculus molossinus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS  BAC end Sequences of Library MSMg01
TITLE    Unpublished
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 754)

```

```

AUTHORS  Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE    Direct Submission
JOURNAL  Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:hattori@gsc.riken.jp, Uri:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT  Clones are derived from the mouse BAC library MSMg01. For BAC
          library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
          Taikuba Institute, Bio Resource Center.
          The Institute of Physical and Chemical Research (RIKEN) 3-1-1
          Koyadai, Tsukuba, 305-0074 Japan
          phone: 81-298-36-9189, fax: 81-298-36-9199
          e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector      : pBACe3.6
R.Site 1    : EcoRI
R.Site 2    : EcoRI.
FEATURES
source      1..754
          /organism="Mus musculus molossinus"
          /mol_type="genomic DNA"
          /sub_species="molossinus"
          /db_xref="taxon:57486"
          /clone="MSMg01-389D22.TJ"
          /sex="male"
          /tissue_type="mixture of kidney and spleen"
          /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match      16.5%; Score 32.6; DB 9; Length 754;
Best Local Similarity 59.9%; Pred. No. 24;
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy  2  GTAACTTACCAATCAAAGACCTTAACCAAGAGCTGTGCTGCTCATTTATGCTGTGG 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  305 GTGCACCTGCTATGCATATACCTTCACAGAGCATTTGCCATGGCATTAAATGTGTAGG 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  62 GGGACATAGGGGAAAAACATATTGCCCTTCTTCTAC 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  245 AAGCAGTAACTGAATCTGATGATGAATTTCTACTTC 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
CV470085/c
LOCUS   CV470085 841 bp mRNA linear EST 01-OCT-2004
DEFINITION
42879.1 Common Scab-Challenged Tubers Solanum tuberosum cDNA clone
42879 5', mRNA sequence.
ACCESSION CV470085
VERSION   CV470085.1 GI:53694861
KEYWORDS EST.
SOURCE   Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 841)
AUTHORS  Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Laque,M., De
          Koeyer,D., Audy,P., Goyer,C., Li,X-Q., Wang-Pruski,G. and Regan,S.
          Generation of ESTs from common scab-challenged potato tubers
          Unpublished (2004)
          Contact: Barry Flinn
          The Canadian Potato Genome Project - BioAtlantech
          921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
          Email: bflinn@bioatlantech.nb.ca
          Seq primer: T3.
FEATURES
source      Location/Qualifiers
          1..841
          /organism="Solanum tuberosum"
          /mol_type="mRNA"
          /cultivar="Shepody"

```

```

/db_xref="taxon:4113"
/clone="42879"
/tissue_type="Tubers"
/lab_host="XL10-Gold"
/clone_lib="Common Scab-Challenged Tubers"
/notes="Vector: pBluescript II SK(+); XR: Site 1: EcoRI; Site 2: XhoI; supplier: Pathogen-challenge series. Tubers from pathogen-free Solanum tuberosum var. Shepody, clone 1756, were inoculated with Streptomyces scabei EF-35, by applying 1 ml of a spore stock (OB650 = 0.7). Tubers were placed in trays of sand and watered from the bottom for the remainder of the experiment. Inoculated tubers were sampled at 2 days, 1 week, 2 weeks and 4 weeks post-infection. Harvested tubers were washed in distilled water, blotted dry, cut into small pieces and frozen in liquid nitrogen. RNA was isolated from each sample collection, pooled and used for cDNA library construction."

Query Match      16.5%; Score 32.6; DB 7; Length 841;
Best Local Similarity 57.3%; Pred. No. 25;
Matches 59; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 79 CATATTGCTTCTTCTACAGAGCGCTTCAGTCGCCATAATCTATGCGGCCCAATTTG 138
Db 802 CATTTCTCTTCTTCTATATTTGACACCACTGTTGGGAATTCCTTTGCCCTGAAGTCC 743

QY 139 GGTTCAGTCTGCTTTCCAGTCATGGGAGAGGTAAGGTTAA 181
Db 742 ACCTTCACTTGTGGAAAGCAATGGATACGGAGCGGTTAA 700

RESULT 12
LOCUS      AQ600889      428 bp      DNA      linear      GSS 10-JUN-1999
DEFINITION HS_5317_A1_B06_SF6E_RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plates=893 Col=11 Row=C, genomic survey sequence.
ACCESSION  AQ600889
VERSION     AQ600889.1 GI:5060883
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 428)
AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE     Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE   99380589
PUBMED    1049764
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.bufo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.bufo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 893 row: C column: 11
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 428.
FEATURES
source
1. .428
Location/Qualifiers

```

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=893 Col=11 Row=C"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

ORIGIN
Query Match      16.4%; Score 32.4; DB 8; Length 428;
Best Local Similarity 64.9%; Pred. No. 25;
Matches 48; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 122 CTTGGCGCCCAATTTGGGTTTCAGTTGCTTCCAGCTATGGGAGAGTAAGGTTAA 181
Db 48 CTTGGCTTCAGAAATTCGCTTTCATTTGCTTCTCATCAATTGAGAAAAGCTTAGGCTAA 107

QY 182 ACCAAACCGTAAAT 195
Db 108 ACCCACAGTAAAT 121

RESULT 13
LOCUS      AQ240675      473 bp      DNA      linear      GSS 30-SEP-1998
DEFINITION CIT-HSP-2386K8.TR.1 CIT-HSP Homo sapiens genomic clone 2386K8,
genomic survey sequence.
ACCESSION  AQ240675
VERSION     AQ240675.1 GI:3672873
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 473)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL   Unpublished (1998)
COMMENT   Other_GSSs: CIT-HSP-2386K8.TF.1
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source
1. .473
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2386K8"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN
Query Match      16.4%; Score 32.4; DB 8; Length 473;
Best Local Similarity 51.4%; Pred. No. 26;

```

```
Matches 75; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy 48 ATTATGCTGTGGGGGACATAGGGGGAACACATATTCCTTCTTACAGAGCGCTTC 107
Db 123 ATCCATAAAGTATGTATCTCTACAGGCAATATATGCGCTCCCTCAACAAGTCAAACTC 182
Qy 108 AGTCGCCATATTAATCTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTTCCTCAGCTATGGGG 167
Db 183 ATGCGCAAAAAGATGAGGGGCACATGTTGGTTTCGCTACGAACATAGTTGGCACGCTC 242
Qy 168 AGAGGTAAGGTTAAACCAACCGTAA 193
Db 243 AGGGTTTGGTTCCAGGATACCAAA 268

RESULT 14
CK405143
LOCUS 945 bp mRNA linear EST 05-JAN-2004
DEFINITION AUF_ifspn_232_f23 Ictalurus furcatus spleen cDNA library Ictalurus
furcatus cDNA 5', mRNA sequence.
ACCESSION CK405143
VERSION CK405143.1 GI:40564502
KEYWORDS EST.
SOURCE Ictalurus furcatus
ORGANISM Ictalurus furcatus
REFERENCE 1 (bases 1 to 945)
AUTHORS Liu, Z., Li, P., Liu, L., He, C., Kucuktas, H., Feng, J., Chen, L.,
Peatman, E., Baopraetukul, P., Simmons, M., Muir, W., Grizzle, J.,
Dunham, R. and Brady, Y.
30,000 new catfish ESTs: new resources for functional analysis of
genes involved in aquaculture performance traits
TITLE 30,000 new catfish ESTs: new resources for functional analysis of
genes involved in aquaculture performance traits
JOURNAL Unpublished (2004)
COMMENT Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: T7.
FEATURES
source
Location/Qualifiers
1..945
/organism="Ictalurus furcatus"
/mol_type="mRNA"
/db_xref="taxon:66913"
/clone_lib="Ictalurus furcatus spleen cDNA library"
/note="Organ: Spleen; Vector: pSport1; Site_1: NotI;
Site_2: SalI"

ORIGIN
Query Match 16.4%; Score 32.4; DB 7; Length 945;
Best Local Similarity 49.1%; Pred. No. 29;
Matches 81; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
Qy 30 AACAGGACTGTCGTGTCATTTATGCTGTGGGGGACATAGGGGGAACACATATTCGCTT 89
Db 681 AAACCTGGCAGTCAGGGCTTCTTCTCGAGGTGGTGCATTCGCGCGGAAAAAATATCGAACT 740
Qy 90 CTTCTACAGAGGCTTCAGTCGCCATAATTTACTTGGCGCCCAATTTTGGGTTTCAGTTG 149
Db 741 GCCACAAATGAGCGTGGTGAAGCCATGAAGATTTTGGACCCCTNNGAGAGGTGNAGAG 800
Qy 150 CTGTTTCCAGCTATCGGGGAGAGGTAAAGGTTAAACCAACCGTAA 194
Db 801 ATCGTATCGGCACTGGAGAAAGAGGAGGAGCGGAGAGAAA 845
```

RESULT 15

```
AI235685 170 bp mRNA linear EST 31-JAN-1999
LOCUS EST232247 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
DEFINITION ROVCV03 3' end, mRNA sequence.
ACCESSION AI235685
VERSION AI235685.1 GI:3829191
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
REFERENCE 1 (bases 1 to 170)
AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.
Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat
Gene Index
TITLE Unpublished (1998)
JOURNAL Contact: Lee, NH
COMMENT The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
FEATURES
source
Location/Qualifiers
1..170
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="ATCC (inhost):2041157"
/db_xref="taxon:10118"
/clone_lib="ROVCV03"
/note="Organ: ovary; Vector: pT73Pac; Site_1: EcoRI;
Site_2: NotI"

ORIGIN
Query Match 16.3%; Score 32.2; DB 1; Length 170;
Best Local Similarity 50.3%; Pred. No. 25;
Matches 79; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
Qy 35 GACTGTGTCGTGTCATTTATGCTGTGGGGGACATAGGGGGAACACATATTCGCTTCTTCT 94
Db 8 GACAGGAGGTTTATTCATGCTGATGGGGTAGGAAGGCAAGGCGTCTTGAGCTATATT 67
Qy 95 ACAAGAGGCTTCAGTCGCCATAATTTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTT 154
Db 68 ATGTGCTTGATGCGCTCTCTTAAGTTTATAGGCAAAATGTTTAAAGGATTTTACACCGGTT 127
Qy 155 TCCAGCTATGGGAGAGGTAAAGGTTAAACCAACCGT 191
Db 128 TATGGATGTTGATGATGTAATCTTACCTCCATCCT 164
```

Search completed: September 26, 2005, 04:37:44
Job time : 1604.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 01:18:09 ; Search time 76.8991 Seconds
(without alignments)
4191.810 Million cell updates/sec

Title: US-09-551-494-4_COPY_1_197

Perfect score: 197

Sequence: 1 agtaaacctaccatcaaaa.....ttaaaccaaacgcgtaaatcg 197

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA.*
- 1: /cgm2_6/ptodata/1/ina/5A.COMB.seq.*
 - 2: /cgm2_6/ptodata/1/ina/5B.COMB.seq.*
 - 3: /cgm2_6/ptodata/1/ina/6A.COMB.seq.*
 - 4: /cgm2_6/ptodata/1/ina/6B.COMB.seq.*
 - 5: /cgm2_6/ptodata/1/ina/PCTUS.COMB.seq.*
 - 6: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31.6	16.0	325034	4	US-09-949-016-14957
2	31.6	16.0	389504	4	US-09-949-016-11774
c 3	30.6	15.5	128723	4	US-09-949-016-17533
c 4	30	15.2	43576	4	US-09-676-519-19
5	29.4	14.9	1989	4	US-09-134-000C-2277
6	29.4	14.9	10401	4	US-08-956-171E-149
7	29.4	14.9	10401	4	US-08-781-986A-149
8	29	14.7	84296	4	US-09-949-016-17375
9	28.6	14.5	3527	4	US-09-949-016-1701
10	28.6	14.5	18679	4	US-09-949-016-13443
11	28.6	14.5	298336	4	US-09-949-016-16600
12	28.2	14.3	1664976	4	US-08-916-421B-1
c 13	28.2	14.3	1664976	4	US-09-692-570-1
14	28	14.2	601	4	US-09-949-016-164388
15	28	14.2	56147	4	US-09-949-016-16352
c 16	27.6	14.0	53336	4	US-09-949-016-12500
c 17	27.6	14.0	53337	4	US-09-949-016-16092
c 18	27.6	14.0	88490	4	US-09-949-016-12758
c 19	27.6	14.0	88736	4	US-09-949-016-14222
20	27.6	14.0	114842	4	US-09-949-016-14993
21	27.4	13.9	400	4	US-08-956-171E-3998
22	27.4	13.9	400	4	US-08-781-986A-3998
23	27.4	13.9	601	4	US-09-949-016-86842
24	27.4	13.9	601	4	US-09-949-016-86843
c 25	27.4	13.9	1266	3	US-08-718-904-75
c 26	27.4	13.9	1266	4	US-09-449-249-75
27	27.4	13.9	2139	4	US-09-543-681A-3975

28	27.4	13.9	247781	4	US-09-949-016-14193
29	27.4	13.9	346112	4	US-09-949-016-13165
c 30	27.4	13.9	451924	4	US-09-949-016-12896
c 31	27.4	13.9	451925	4	US-09-949-016-17305
32	27.4	13.9	818128	4	US-09-949-016-14546
33	27.4	13.9	818128	4	US-09-949-016-14547
34	27.4	13.9	818128	4	US-09-949-016-14548
35	27.4	13.9	818128	4	US-09-949-016-14549
36	27.4	13.9	818128	4	US-09-949-016-14550
37	27.4	13.9	818128	4	US-09-949-016-14551
38	27.4	13.9	818128	4	US-09-949-016-14552
39	27.4	13.9	818128	4	US-09-949-016-14553
40	27.4	13.9	818128	4	US-09-949-016-14554
41	27.4	13.9	818128	4	US-09-949-016-14555
42	27.4	13.9	818128	4	US-09-949-016-14556
43	27.4	13.9	818128	4	US-09-949-016-14557
44	27.4	13.9	818128	4	US-09-949-016-14558
45	27.4	13.9	818128	4	US-09-949-016-14559

ALIGNMENTS

RESULT 1

US-09-949-016-14957
; Sequence 14957, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14957
; LENGTH: 325034
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14957

Query Match 16.0%; Score 31.6; DB 4; Length 325034;
Best Local Similarity 58.5%; Pred. No. 4.4;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 2 GTAACTTACCATCAAGACCTAACCAAGACTGTGTCATTTATGCTGTTGG 61
Db 261036 GTAACTTATATATCTTAAATTTATCTGGGAATTTTGAAGTCGGTGAAGTTGG 261095

Qy 62 GGCACATAGGGGAAACATATTGCCCTTCTCTA 95
Db 261096 GGTCTCAGGAGGAGATCAATGTTACCCCTTCCA 261129

RESULT 2

US-09-949-016-11774
; Sequence 11774, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11774
; LENGTH: 389504
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11774

Query Match          16.0%; Score 31.6; DB 4; Length 389504;
Best Local Similarity 58.5%; Pred. No. 4.8; Mismatches 39; Indels 0; Gaps 0;
Matches 55; Conservative 0;

Qy  2  GTAACATTACCAATCAAAAGACCTTAACCAACAGGACTGTGTGTCTATTATGCTGTGG 61
Db  277506  GTAACTTATATATCTTAAAATTTATCTGTGGAAATTTTGAAGTCTGGTGAAGTTGG 277565

Qy  62  GGCACATAGAGGGGAAACATATTCCTTCTTA 95
Db  277566  GGTCTCAGGAGGAATCATGTTTACCCCTTCCA 277599

RESULT 3
US-09-949-016-17533/c
; Sequence 17533, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17533
; LENGTH: 128723
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17533

Query Match          15.5%; Score 30.6; DB 4; Length 128723;
Best Local Similarity 52.8%; Pred. No. 6.3; Mismatches 59; Indels 0; Gaps 0;
Matches 66; Conservative 0;

Qy  72  GGGAAACATATTCCTTCTTCTACAGAGCCCTTCAGTCGCCATATTAATTACTTGGGCC 131
Db  16352  GAGTACTCATTCAGTCTGCTCTCTTAACAGACCTTTAAAGAGCTCATTTTGTGTTT 16293

Qy  132  AATTTGGGTTTCAGTGTCTGTTTCCAGCTATGGGAGAGTAAGGTTAAACCAACCGT 191
Db  16292  GTTTTTGGTCAGAGATCTTTTCCCTTACTGAAACTGAGAGTATTAACCAAAAC 16233

Qy  192  AATC 196
Db  16232  AATC 16228

RESULT 4
US-09-676-519-19/c
; Sequence 19, Application US/09676519
; Patent No. 6737508
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
```

```
; APPLICANT: PELLETIER, JERRY
; APPLICANT: GROS, PHILIPPE
; APPLICANT: DUBOW, MICHAEL
; TITLE OF INVENTION: DNA SEQUENCES FROM STAPHYLOCOCCUS AUREUS BACTERIOPHAGES
; FILE REFERENCE: 3A, 77, AND 96 THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES
; CURRENT APPLICATION NUMBER: US/09/676,519
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/407,804
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 60/110,992
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 43576
; TYPE: DNA
; ORGANISM: Staphylococcus bacteriophage
US-09-676-519-19

Query Match          15.2%; Score 30; DB 4; Length 43576;
Best Local Similarity 55.9%; Pred. No. 6; Mismatches 45; Indels 0; Gaps 0;
Matches 57; Conservative 0;

Qy  78  ACATATTGCTTCTTCTACAGAGCCCTTCAGTCGCCATAAATTACTTGGGCCCAATTTT 137
Db  8895  AAATTCTGCCAACTTATTCATGAAGTTATCTAGGCTATTTTCTTGTCTGACATATTC 8836

Qy  138  GGTTTTCAGTTGTGTTTCCAGCTATGGGAGAGGTAAGTT 179
Db  8835  GGGTCACTCTGCATTTTCCATTTGGTGTAGCGGTTTCGTT 8794

RESULT 5
US-09-134-000C-2277
; Sequence 2277, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2277
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2277

Query Match          14.9%; Score 29.4; DB 4; Length 1989;
Best Local Similarity 58.6%; Pred. No. 1.9; Mismatches 36; Indels 0; Gaps 0;
Matches 51; Conservative 0;

Qy  67  ATAGGGGAAAACATATATGCTTCTTCTACAAGAGCGCTTCAGTCGCCATAAATTACTTGG 126
Db  1216  AGAGGCATGAACGTTTTTGGCGACTTTTCTTGATTTTATCATTTTCAACAATATTATTGG 1275

Qy  127  CGCCCAATTTGGTTTCAGTTGCTGT 153
Db  1276  GGCTTCATTTACGGTTTCGTTTTCGGT 1302

RESULT 6
US-08-956-171E-149
; Sequence 149, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
```



```
;
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
;
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 149:
US-08-956-171E-149

Query Match 14.9%; Score 29.4; DB 4; Length 10401;
Best Local Similarity 55.3%; Pred. No. 4.6;
Matches 57; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 78 ACATATTGCCCTTCTTCTACAGAGCCCTTCAGTCGCCCATATTACTTGGCGCCCAATTTT 137
Db 6710 AAATTCGCCCACTTATTCATGAACCTATCTAGCGCTATTTTCTTTGCTGACATATTC 6769

Qy 138 GGGTTTCAGTTCGTTTCCAGCTATCGGAGAGGTAAGTTA 180
Db 6770 GGGGTCATCTGCATTTTCCATTTGCTGTGAGCGTTTCGTTA 6812

RESULT 7
US-08-781-986A-149
; Sequence 149, Application US/08/781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
;
; Query Match 14.7%; Score 29; DB 4; Length 84296;
; Best Local Similarity 61.0%; Pred. No. 20;
; Matches 47; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
; Qy 30 AACAGGACTGTCGTGCTCAATTATGCTGTGGGGGACATAGGGGAAAAACATATTGCCTT 89
```

```
;
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
;
; INFORMATION FOR SEQ ID NO: 149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-781-986A-149
Query Match 14.9%; Score 29.4; DB 4; Length 10401;
Best Local Similarity 55.3%; Pred. No. 4.6;
Matches 57; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 78 ACATATTGCCCTTCTTCTACAGAGCCCTTCAGTCGCCCATATTACTTGGCGCCCAATTTT 137
Db 6710 AAATTCGCCCACTTATTCATGAACCTATCTAGCGCTATTTTCTTTGCTGACATATTC 6769

Qy 138 GGGTTTCAGTTCGTTTCCAGCTATCGGAGAGGTAAGTTA 180
Db 6770 GGGGTCATCTGCATTTTCCATTTGCTGTGAGCGTTTCGTTA 6812

RESULT 8
US-09-949-016-17375
; Sequence 17375, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17375
; LENGTH: 84296
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84296)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17375

Query Match 14.7%; Score 29; DB 4; Length 84296;
Best Local Similarity 61.0%; Pred. No. 20;
Matches 47; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
; Qy 30 AACAGGACTGTCGTGCTCAATTATGCTGTGGGGGACATAGGGGAAAAACATATTGCCTT 89
```

Db 65035 AGCAAGCTGCTCGAACAGTGCTTGGGGTGGAGTAGGGTCAATGGTATCCCTT 65094
Qy 90 CTTCTACAAGGCGCTT 106
Db 65095 CTTTATAAGAGCTCTT 65111

RESULT 9
US-09-949-016-1701
; Sequence 1701, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1701
; LENGTH: 3527
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1701

Query Match 14.5%; Score 28.6; DB 4; Length 3527;
Best Local Similarity 53.0%; Pred. No. 5.2;
Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 4 AAACCTTACCAATCAAAAGACCTAACCAACAGACTGCTGCTGCTATTTATGCTCTTGGG 63
Db 2367 ATACACAGCCCAAGTAAAGTTTGAACAAAAAATGTAGAAGCTATTGGACTTCTGGAG 2426
Qy 64 GACATAGGGGAAAAACATATTGCTTCTTACAAAGAGGCTTCAGTCGCCATAA 118
Db 2427 GACAAAGAGCAAGCAGATTATGAGCTATACAAAGAGCTCTTAATCCTGATAA 2481

RESULT 10
US-09-949-016-13443
; Sequence 13443, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13443
; LENGTH: 18679
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13443
Query Match 14.5%; Score 28.6; DB 4; Length 18679;
Best Local Similarity 53.0%; Pred. No. 13;

Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 4 AAACCTTACCAATCAAAAGACCTAACCAACAGACTGCTGCTGCTATTTATGCTCTTGGG 63
Db 15519 ATACACAGCCCAAGTAAAGTTTGAACAAAAAATGTAGAAGCTATTGGACTTCTGGAG 15578
Qy 64 GACATAGGGGAAAAACATATTGCTTCTTCTACAAAGAGGCTTCAGTCGCCATAA 118
Db 15579 GACAAAGAGCAAGCAGATTATGAGCTATACAAAGCCTCTAATCCTGATAA 15633

RESULT 11
US-09-949-016-16600
; Sequence 16600, Application US/09949016
; Patent No. 8812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16600
; LENGTH: 298336
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(298336)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16600

Query Match 14.5%; Score 28.6; DB 4; Length 298336;
Best Local Similarity 54.2%; Pred. No. 57;
Matches 58; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
Qy 25 TAACCAACAGAGCTGCTGCTCATTTATGCTGTGGGGACATAGGGGAAACATATT 84
Db 164133 TCAGCAATAATACTGTTTATGTTTCTTATCATTTGTGTGTTACCGGAGTAGCATTTTT 164192
Qy 85 GCCTTCTTCTACAAAGAGGCTTCAGTCGCCATAATTACTTGGCGCCC 131
Db 164193 AAATTTCTTCAAGAACCTTTCATTTTCATTCGCAACTTGGCTCAC 164239

RESULT 12
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103398)..(103398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
```

```
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1
```

Query Match 14.3%; Score 28.2; DB 4; Length 1664976;

Best Local Similarity 54.3%; Pred. No. 1.8e+02;

Matches 57; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```
Qy 75 AAAACATATTGCTCTTCTTACAAGAGGCCCTTCAGTCGCCATAATTACTTGGCGCCCAAT 134
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 273184 AAAAAATAGTGACCTTGTTATTATAGGCCCTCAAAATCCAATACTTCCATAGGTCCAAT 273125
```

```
Qy 135 TTTCGGTTTCAGTGTCTTTCACGACTATGGGGAGAGGTAAGGTT 179
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 273124 TTTAAGTTTAAATGGAATTAAGAGACTATTAAAGATATAAAAGGT 273080
```

RESULT 13

```
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (319228)..(319228)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
```

```
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)

; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

Query Match 14.3%; Score 28.2; DB 4; Length 1664976;
Best Local Similarity 54.3%; Pred. No. 1.8e+02;
Matches 57; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 75 AAAACATATTGCTTCTTCTACAAGAGCGCTTCAGTCGCCATAATTAATTGCGCGCCCAAT 134
|||||
Db 273184 AAAAATAGTGACCTTGTTATTATAGGCGCCTCAAAATCCAATAACTTCCATAGGTCCAAT 273125
|||||

Qy 135 TTTCGGTTTCAGTTGCTGTTTCCAGCTATGCGGAGAGTAAGTTT 179
|||||
Db 273124 TTTAAGTTTAAATGGAATTAAGAGCTATTAAAGATATAAAAGGT 273080
|||||

RESULT 14
US-09-949-016-164388
; Sequence 164388, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164388
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-164388

Query Match 14.2%; Score 28; DB 4; Length 601;
Best Local Similarity 63.2%; Pred. No. 3.4;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 34 GGACTGCTGGTTCATTATGCTGTTGGGGACATAGGGGAAACATATTGCTTCTTC 93
|||||
Db 132 GGAGTATGTTTCTCTTTTGTCAATAGTGGAAATAAATAATGCAATTTTC 191
|||||

Qy 94 TACAAGAG 101
|||||
Db 192 CAGAAGAG 199

RESULT 15
US-09-949-016-16352
; Sequence 16352, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16352
; LENGTH: 56147
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16352

Query Match      14.2%; Score 28; DB 4; Length 56147;
Best Local Similarity 63.2%; Pred. No. 39;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy      34 GGACTGTGCTGCTGCTTATCTGTTGGGGACATAGGGGAAACATATTGCCTTCTTC 93
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      5082 GGAGTATGTTTTTCTTTTTCATTTGTCAGAAATAGTGGAAATAAATATGCATTTTGC 5141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy      94 TACAAGAG 101
      ||| ||| |||
Db      5142 CAGAAGAG 5149
```

Search completed: September 26, 2005, 04:42:29
Job time : 90.8991 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 01:30:34 ; Search time 347.342 Seconds
(without alignments)
3792.150 Million cell updates/sec

Title: US-09-551-494-4_COPY_1_197
Perfect score: 197
Sequence: 1 agtaaaattaccatcaaaa.....ttaaaccaaacgcgtaaatcg 197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
 - 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
 - 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
 - 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
 - 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
 - 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
 - 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 - 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
 - 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
 - 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
 - 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	197	100.0	1058	15	US-10-321-434-2
2	197	100.0	1058	19	US-10-315-826-4
3	184.2	93.5	1059	15	US-10-321-434-1
4	34.8	17.7	734	20	US-10-425-115-41144
5	34.4	17.5	3673778	16	US-10-312-841-1
6	33.2	16.9	198285	9	US-09-880-107-3814
7	33.2	16.9	198285	19	US-10-775-169-338

c	8	33.2	16.9	198285	22	US-10-756-149-12	Sequence 12, Appl
c	9	31.6	16.0	2140405	17	US-10-027-633-76212	Sequence 76212, A
c	10	31.6	16.0	2140405	13	US-10-027-633-76212	Sequence 76212, A
c	11	30.6	15.4	144723	13	US-10-087-192-1576	Sequence 1576, Ap
c	12	30.4	15.4	819	14	US-10-198-846-6058	Sequence 6058, Ap
c	13	30.4	15.4	1196	14	US-10-198-846-5470	Sequence 5470, Ap
c	14	30	15.2	122814	21	US-10-741-600-17652	Sequence 17652, A
c	15	30	15.2	249487	13	US-10-026-188-3	Sequence 3, Appli
c	16	30	15.2	1691139	14	US-10-067-514-1	Sequence 1, Appli
c	17	30	15.2	1691139	17	US-10-419-723-1	Sequence 1, Appli
c	18	29.8	15.1	642	13	US-10-027-632-189483	Sequence 189483,
c	19	29.8	15.1	642	17	US-10-027-632-189483	Sequence 189483,
c	20	29.6	15.0	506	13	US-10-027-632-44151	Sequence 44151, A
c	21	29.6	15.0	506	17	US-10-027-632-44151	Sequence 44151, A
c	22	29.6	15.0	622	20	US-10-357-930-52336	Sequence 52336, A
c	23	29.4	14.9	518	10	US-09-814-353-1929	Sequence 1929, Ap
c	24	29.4	14.9	518	10	US-09-814-353-8276	Sequence 8276, Ap
c	25	29.4	14.9	8726	9	US-09-070-927A-95	Sequence 95, Appl
c	26	29.4	14.9	10401	8	US-08-781-986A-149	Sequence 149, App
c	27	29.4	14.9	10401	18	US-10-329-624-149	Sequence 149, App
c	28	29.4	14.9	1691140	22	US-10-868-397-1	Sequence 1, Appli
c	29	29.2	14.8	389	18	US-10-424-599-119662	Sequence 119662,
c	30	29.2	14.8	701	13	US-10-027-632-15535	Sequence 15535, A
c	31	29.2	14.8	701	13	US-10-027-632-15535	Sequence 15535, A
c	32	29.2	14.8	701	13	US-10-027-632-15536	Sequence 15536, A
c	33	29.2	14.8	701	17	US-10-027-632-145843	Sequence 145843,
c	34	29.2	14.8	701	17	US-10-027-632-15535	Sequence 15535, A
c	35	29.2	14.8	701	17	US-10-027-632-15536	Sequence 15536, A
c	36	29.2	14.8	33271	20	US-10-027-632-145843	Sequence 145843,
c	37	29.2	14.8	158811	20	US-10-719-993-6996	Sequence 6996, Ap
c	38	29.2	14.8	786452	20	US-10-723-860-2720	Sequence 2720, Ap
c	39	29	14.7	547	13	US-10-027-632-195673	Sequence 195673,
c	40	29	14.7	547	13	US-10-027-632-195674	Sequence 195674,
c	41	29	14.7	547	17	US-10-027-632-195673	Sequence 195673,
c	42	29	14.7	547	17	US-10-027-632-195674	Sequence 195674,
c	43	29	14.7	600	22	US-10-972-079-11384	Sequence 11384, A
c	44	29	14.7	3283	13	US-10-027-632-114761	Sequence 114761,
c	45	29	14.7	3283	17	US-10-027-632-114761	Sequence 114761,

ALIGNMENTS

RESULT 1

US-10-321-434-2
; Sequence 2, Application US/10321434
; Publication No. US20030135882A1
; GENERAL INFORMATION:
; APPLICANT: Metzlaiff, Michael
; APPLICANT: Meulewater, Frank
; APPLICANT: Gossel, Veronique
; APPLICANT: Fach, Ina
; TITLE OF INVENTION: Improved methods and means for delivering inhibitory RNA to plants
; TITLE OF INVENTION: applications thereof
; FILE REFERENCE: FROMOD
; CURRENT APPLICATION NUMBER: US/10321434
; CURRENT FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1058
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cdna sequence of the genome of satellite tobacco mosaic virus
US-10-321-434-2

Query Match 100.0%; Score 197; DB 15; Length 1058;

Best Local Similarity 100.0%; Pred. No. 3.8e+60;

Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGTAACTTACCATCAAGACCTACACAGACATGCTGCTATTATGCTGTG 60

|||||

Db 1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTGCTGTCATTTATGCTGTG 60
QY 61 GGGACATAGGGGAAACATATTGCTTCTTACAGAGGCTTCAGTCGCCATAATT 120
Db 61 GGGACATAGGGGAAACATATTGCTTCTTACAGAGGCTTCAGTCGCCATAATT 120
QY 121 ACTTGGGGCCCAATTTTGGGTTTCAGTTGCTGTTTCCAGCTATGCGGAGAGGTAAGGTTA 180
Db 121 ACTTGGGGCCCAATTTTGGGTTTCAGTTGCTGTTTCCAGCTATGCGGAGAGGTAAGGTTA 180
QY 181 AACCAAAACCGTAAATCG 197
Db 181 AACCAAAACCGTAAATCG 197

RESULT 2
US-10-315-826-4
; Sequence 4, Application US/10315826
; Publication No. US20040110141A1
; GENERAL INFORMATION:
; APPLICANT: Pusey, Marc L.
; APPLICANT: Dowell, Jeffrey
; APPLICANT: Ng, Joseph D.
; TITLE OF INVENTION: Nucleic Acid Detector and Method of
; TITLE OF INVENTION: Detecting the Presence of Nucleic Acid Targets Within a
; TITLE OF INVENTION: Biological Sample
; FILE REFERENCE: 38470/249510
; CURRENT APPLICATION NUMBER: US/10/315,826
; CURRENT FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1058
; TYPE: DNA
; ORGANISM: Satellite Tobacco Mosaic Virus
US-10-315-826-4

Query Match 100.0%; Score 197; DB 19; Length 1058;
Best Local Similarity 100.0%; Pred. No. 3.9e-60;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTGCTGTCATTTATGCTGTG 60
Db 1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTGCTGTCATTTATGCTGTG 60
QY 61 GGGACATAGGGGAAACATATTGCTTCTTACAGAGGCTTCAGTCGCCATAATT 120
Db 61 GGGACATAGGGGAAACATATTGCTTCTTACAGAGGCTTCAGTCGCCATAATT 120
QY 121 ACTTGGGGCCCAATTTTGGGTTTCAGTTGCTGTTTCCAGCTATGCGGAGAGGTAAGGTTA 180
Db 121 ACTTGGGGCCCAATTTTGGGTTTCAGTTGCTGTTTCCAGCTATGCGGAGAGGTAAGGTTA 180
QY 181 AACCAAAACCGTAAATCG 197
Db 181 AACCAAAACCGTAAATCG 197

RESULT 3
US-10-321-434-1
; Sequence 1, Application US/10321434
; Publication No. US20030135882A1
; GENERAL INFORMATION:
; APPLICANT: Metzlaff, Michael
; APPLICANT: Meulwater, Frank
; APPLICANT: Gossel, Veronique
; APPLICANT: Fach, Ina
; TITLE OF INVENTION: Improved methods and means for delivering inhibitory RNA to plant
; TITLE OF INVENTION: applications thereof
; FILE REFERENCE: FKMOD
; CURRENT APPLICATION NUMBER: US/10/321,434
; CURRENT FILING DATE: 2002-12-18

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA sequence of the genome of satellite tobacco mosaic virus clo
; OTHER INFORMATION: ned into pSTMV-1
US-10-321-434-1

Query Match 93.5%; Score 184.2; DB 15; Length 1059;
Best Local Similarity 95.9%; Pred. No. 1.6e-55;
Matches 189; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 AGTAACTTACCAATCAAAAGACCTAACCAACAGGACTGTGCTGTCATTTATGCTGTG 60
Db 1 AGTAAACTTTACCAATCAAAAGACCTAACCAACAGGACTGTGCTGTCATTTATGCTGTG 60
QY 61 GGGACATAGGGGAAACATATTGCTTCTTACAGAGGCTTCAGTCGCCATAATT 120
Db 61 GGGACATAGGGGAAACATATTGCTTCTTACAGAGGCTTCAGTCGCCATAATT 120
QY 121 ACTTGGGGCCCAATTTTGGGTTTCAGTTGCTGTTTCCAGCTATGCGGAGAGGTAAGGTTA 180
Db 121 ACTTGGGGCCCAATTTTGGGTTTCAGTTGCTGTTTCCAGCTATGCGGAGAGGTAAGGTTA 180
QY 181 AACCAAAACCGTAAATCG 197
Db 181 AACCAAAACCGTAAATCG 197

RESULT 4
US-10-425-115-41144
; Sequence 41144, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 41144
; LENGTH: 734
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_137528C.1
US-10-425-115-41144

Query Match 17.7%; Score 34.8; DB 20; Length 734;
Best Local Similarity 52.0%; Pred. No. 0.11; 72; Indels 0; Gaps 0;
Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 41 CGTGGTCATTTATGCTGTGCGGACATAGGGGAAACATATTGCTTCTTCTACAGA 100
Db 342 CTGACCATTTGCGATTTTCAGGAGCACTGGGAAAAAAAATTCGCTGTATTACCTGA 401
QY 101 GGCCTTCAGTCGCCATAATTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTTTCCAGC 160
Db 402 CTCATTGGGTGGCTGCGACCTGCTGGGAGCAACACGATGCTGCACATTTTCAGGCCTTCACC 461
QY 161 TATGGGAGAGGTAAGGTTAAACCAACCG 190
Db 462 TTTATCTGGAGAAAGAGATAAAATTTGTACCG 491
RESULT 5

US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1
Query Match 17.5%; Score 34.4; DB 16; Length 3673778;
Best Local Similarity 52.0%; Pred. No. 5.1;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy 35 GACTGTCGTGTCATTTATGCTGTGGGGACATAGGGGAAAAACATATTGCCCTTCTTCT 94
Db 3020478 GATTCTCTGGGTATTGGGAAGATTTTGGGGATAAATGTTAAGAGATATTAAATTTT 3020537
Qy 95 ACAAGAGCCCTTCAGTCGCCATAATTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTT 154
Db 3020538 ATTAGTATTTTGAATTTATGTTTGTGGGTAGTTATATGTAATATTAGATCGTTT 3020597
Qy 155 TCCAGCTATGGGGAGAGTAAGGTTAA 182
Db 3020598 TTTATTGTTAGGAGATTAAAGGTTAA 3020625
RESULT 6
US-09-880-107-3814/c
; Sequence 3814, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3814
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X87344
US-09-880-107-3814
Query Match 16.9%; Score 33.2; DB 9; Length 198285;
Best Local Similarity 51.3%; Pred. No. 4.2;
Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 35 GACTGTCGTGTCATTTATGCTGTGGGGACATAGGGGAAAAACATATTGCCCTTCTTCT 94
Db 162395 GATTCTCTGGGTATTGGGAAGATCCTTGGGGACAAATGCCAAGAAGACACCAACTTCTTT 162336

Qy 95 ACAAGAGCCCTTCAGTCGCCATAATTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTT 154
Db 162335 ATCAGCATCTCTCAGAAATCCATGCCCTGTGGGTAGTTATATGTAATATCAGAAATCGCCC 162276
Qy 155 TCCAGCTATGGGGAGAGTAAGGTTAAACC 184
Db 162275 TCCACTGCCAGGGAGACCCCAAGGCTTAATC 162246
RESULT 7
US-10-775-169-338/c
; Sequence 338, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 338
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-338
Query Match 16.9%; Score 33.2; DB 19; Length 198285;
Best Local Similarity 51.3%; Pred. No. 4.2;
Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 35 GACTGTCGTGTCATTTATGCTGTGGGGACATAGGGGAAAAACATATTGCCCTTCTTCT 94
Db 162395 GATTCTCTGGGTATTGGGAAGATCCTTGGGGACAAATGCCAAGAAGACACCAACTTCTTT 162336
Qy 95 ACAAGAGCCCTTCAGTCGCCATAATTACTTGGCGCCCAATTTTGGGTTTCAGTTGCTGTT 154
Db 162335 ATCAGCATCTCTCAGAAATCCATGCCCTGTGGGTAGTTATATGTAATATCAGAAATCGCCC 162276
Qy 155 TCCAGCTATGGGGAGAGTAAGGTTAAACC 184
Db 162275 TCCACTGCCAGGGAGACCCCAAGGCTTAATC 162246

RESULT 8
US-10-756-149-12/c
; Sequence 12, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-12
Query Match 16.9%; Score 33.2; DB 22; Length 198285;
Best Local Similarity 51.3%; Pred. No. 4.2;
Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 35 GACTGTCGTGTCATTTATGCTGTGGGGACATAGGGGAAAAACATATTGCCCTTCTTCT 94

```
Db 162395 GATTCTCTGGTATTTGGAGATCTTGGGACAAATCCCAAGAACACCAACTCTCTTT 162336
Qy 95 ACAAGAGCCTTCAAGTCCGCAATAATTACTTGGCCGCCAAATTTGGGTTTCAAGTTCGCTGTT 154
Db 162335 ATCAGCATCTCTCAGAATCCATCCCTGCTGGGTAGTTATATGTGAATATCAGAATCGCCC 162276
Qy 155 TCCAGCTATGGGAGAGTAGGTTAAACC 184
Db 162275 TCCATGCCAGGAGACCAAGGCTTAAATC 162246
```

```
RESULT 9
US-10-027-632-76212/c
; Sequence 76212, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76212
; LENGTH: 2140405
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2140405)
; OTHER INFORMATION: n = A,T,C or G
```

```
US-10-027-632-76212
Query Match 16.0%; Score 31.6; DB 13; Length 2140405;
Best Local Similarity 50.7%; Pred. No. 42;
Matches 73; Conservative 1; Mismatches 70; Indels 0; Gaps 0;

Qy 50 TTATGCTGTTGGGGACATAGGGGAAACATATTGCTTCTTCAAGAGGCTTCAG 109
Db 680699 TTATCGAGTTGGCAACATGTGGGAGCTCAAGAAAACGATCAGGAGCACTGCAC 680640

Qy 110 TCGCCATAATTACTTGGCGCCCAATTTGGGTTTCAGTTGCTTCCAGCTATGGGAG 169
Db 680639 ACCTTAAATATGCTTGAAGCCAAATGWWGGGCTGAAACAAACCCCTTACAGTCTGTGAGAAA 680580

Qy 170 AGGTAAGGTTAAACCAACCGTAA 193
Db 680579 TGAATAATCTTAGAGCAAGCAATAA 680556

RESULT 10
US-10-027-632-76212/c
; Sequence 76212, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
Query Match 16.0%; Score 31.6; DB 13; Length 2140405;
Best Local Similarity 50.7%; Pred. No. 42;
Matches 73; Conservative 1; Mismatches 70; Indels 0; Gaps 0;

Qy 50 TTATGCTGTTGGGGACATAGGGGAAACATATTGCTTCTTCAAGAGGCTTCAG 109
Db 680699 TTATCGAGTTGGCAACATGTGGGAGCTCAAGAAAACGATCAGGAGCACTGCAC 680640

Qy 110 TCGCCATAATTACTTGGCGCCCAATTTGGGTTTCAGTTGCTTCCAGCTATGGGAG 169
Db 680639 ACCTTAAATATGCTTGAAGCCAAATGWWGGGCTGAAACAAACCCCTTACAGTCTGTGAGAAA 680580

Qy 170 AGGTAAGGTTAAACCAACCGTAA 193
Db 680579 TGAATAATCTTAGAGCAAGCAATAA 680556
```

```
RESULT 10
US-10-027-632-76212/c
; Sequence 76212, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76212
; LENGTH: 2140405
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2140405)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-76212
```

```
Query Match 16.0%; Score 31.6; DB 17; Length 2140405;
Best Local Similarity 50.7%; Pred. No. 42;
Matches 73; Conservative 1; Mismatches 70; Indels 0; Gaps 0;

Qy 50 TTATGCTGTTGGGGACATAGGGGAAACATATTGCTTCTTCAAGAGGCTTCAG 109
Db 680699 TTATCGAGTTGGCAACATGTGGGAGCTCAAGAAAACGATCAGGAGCACTGCAC 680640

Qy 110 TCGCCATAATTACTTGGCGCCCAATTTGGGTTTCAGTTGCTTCCAGCTATGGGAG 169
Db 680639 ACCTTAAATATGCTTGAAGCCAAATGWWGGGCTGAAACAAACCCCTTACAGTCTGTGAGAAA 680580

Qy 170 AGGTAAGGTTAAACCAACCGTAA 193
Db 680579 TGAATAATCTTAGAGCAAGCAATAA 680556
```

```
RESULT 11
US-10-087-192-1576/c
; Sequence 1576, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Engelhart, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1576
; LENGTH: 144723
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1576

Query Match 15.5%; Score 30.6; DB 13; Length 144723;
Best Local Similarity 52.8%; Pred. No. 32;
```

```
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
Qy 72 GGGAAAACATATTGCTTCTCTACAGAGGCGCTTCAGTCGCCATAAATTACTTGGCGCC 131
Db 24352 GAGTACTCAATTCAGTCTGCTCTGTAAACAGACCTTTAAAGCTCAATTTTGTGCTTT 24293
Qy 132 AATTTGGTTTCAGTCTGCTTCTCCAGCTATGGGAGAGTAGGTAAACCAACCGT 191
Db 24292 GTTTTGGTCAGAGATCTTTTTCCTTACTGAAACTGAGAGTATTAAACCAAAACAC 24233
Qy 192 AAATC 196
Db 24232 AAATC 24228

RESULT 12
US-10-198-846-6058/c
; Sequence 6058, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yongyao
; APPLICANT: Lillie, James
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6058
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 72, 82, 87, 92, 97, 109, 110, 122, 135, 136, 143, 147,
; LOCATION: 158, 159, 163, 166, 184, 187, 193, 209, 210, 258, 262, 263,
; LOCATION: 264, 276, 332, 355, 359, 374, 381, 387, 388, 390, 391, 392,
; LOCATION: 399, 406, 407, 417, 418, 419, 421, 424, 429, 435, 446
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 471, 478, 502, 507, 508, 510, 543, 547, 583, 584, 585, 586,
; LOCATION: 587, 589, 591, 593, 594, 595, 596, 597, 598, 599, 600, 601,
; LOCATION: 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613,
; LOCATION: 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636,
; LOCATION: 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648,
; LOCATION: 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660,
; LOCATION: 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683,
; LOCATION: 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695,
; LOCATION: 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707,
; LOCATION: 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730,
; LOCATION: 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742,
; LOCATION: 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754,
; LOCATION: 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765
; OTHER INFORMATION: n = A,T,C or G
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777,
; LOCATION: 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789,
; LOCATION: 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801,
; LOCATION: 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812-819
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-6058
Query Match 15.4%; Score 30.4; DB 14; Length 819;
Best Local Similarity 50.5%; Pred. No. 4.5;
Matches 55; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 73 GGAACAATATTGCTTCTCTACAGAGGCGCTTCAGTCGCCATAAATTACTTGGCGCCA 132
Db 413 GGACCAANTATTGCNATGATTNNNNAGAGCNCATTTTNCCTTCATTAAAGGGNACANA 354
Qy 133 ATTTTGGTTTCAGTCTGCTGTTTCCAGCTATGGGAGAGGTAAAGTTAA 181
Db 353 ATCATGGGTACCGGGCCCCCCTAGAGGTAGCGGTATCGATAAGCTTAA 305

RESULT 13
US-10-198-846-5470/c
; Sequence 5470, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5470
; LENGTH: 1196
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 4, 8, 9, 10, 11, 13, 14, 15, 16, 19, 23, 24, 27, 28, 33,
; LOCATION: 37, 38, 40, 41, 44, 45, 46, 48, 49, 52, 54, 56, 58, 59,
; LOCATION: 60, 61, 62, 66, 69, 72, 74, 76, 85, 86, 89, 90, 91, 92,
; LOCATION: 94, 97, 108, 116, 126, 128, 129, 133, 135, 139, 151, 152,
; LOCATION: 154
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 155, 163, 174, 190, 194, 219, 293, 313, 403, 407, 433, 494,
; LOCATION: 499, 508, 509, 515, 516, 517, 521, 551, 557, 569, 604, 667,
; LOCATION: 671, 699, 708, 715, 724, 725, 766, 768, 769, 770, 785, 793,
; LOCATION: 822, 846, 853, 857, 858, 860, 879, 887, 891, 892, 895
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 902, 913, 919, 935, 946, 949, 965, 1002, 1010, 1013, 1022,
; LOCATION: 1023, 1032, 1041, 1060, 1065, 1069, 1071, 1093, 1112, 1118,
; LOCATION: 1120, 1137, 1140, 1150, 1175, 1184, 1185, 1187, 1190, 1192,
; LOCATION: 1193, 1195, 1196
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5470
Query Match 15.4%; Score 30.4; DB 14; Length 1196;
Best Local Similarity 59.7%; Pred. No. 5.3;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

QY 86 CTTCTCTTCAAGAGGCTTCACTGCGCCATAATTACTTGGCGCCAAATTTTGGTTTCA 145
|||||
Db 1130 CCITTTTANAGGGTTTCCTGGGCGCCCTTTTGTGTTGGGCGCCCTTTTGGGTTTCCN 1071
|||||
QY 146 GTTGCTCTTTCC 157
|||||
Db 1070 ANCATNGTTTNC 1059
|||||

RESULT 14
US-10-741-600-17652
; Sequence 17652, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17652
; LENGTH: 122814
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(122814)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17652

Query Match 15.2%; Score 30; DB 21; Length 122814;
Best Local Similarity 55.9%; Pred. No. 49;
Matches 57; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY 15 TCAAAAGACCTTAACCAACAGGACTGTGCTGTCATTATGCTGTTGGGGACATAGGGGG 74
|||||
Db 44412 TCTCAAAGCTTTAGCAGAGTAGCTTTAATTAATGCTGTGGAGGATAGGGGG 44471
|||||
QY 75 AAAACATATTCCTCTTCTTACAAGAGGCTTCAAGTCCCAT 116
|||||
Db 44472 AAAAAGAGACTAGCTTCTTAATAAGAACTTGAAGTCCAT 44513
|||||

RESULT 15
US-10-026-188-3/c
; Sequence 3, Application US/10026188
; Publication No. US20020164645A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Zhang, Yifeng
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
; TITLE OF INVENTION: Ion Channel
; FILE REFERENCE: 02307E-114910US
; CURRENT APPLICATION NUMBER: US/10/026,188
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/259,379
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 249487
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse genomic region containing ltrpc5
US-10-026-188-3

Query Match 15.2%; Score 30; DB 13; Length 249487;
Best Local Similarity 55.9%; Pred. No. 66;
Matches 57; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 2 GTAAACTTACCAATCAAAAGACCTTAACCAACAGGACTGTGCTGCTCATTTATGCTGTTGG 61
|||||
Db 220223 GTCATCTCAGCACTCAGAAGATGGAAGCAAGGTCACTCTTGCTACTTAATGAGTTCA 220164
|||||
QY 62 GGGACATAGGGGGAAACATATTGGCCTTCTTCTACAGAGGC 103
|||||
Db 220163 AGGCCAGCCTGAGATATGTTAGTTCTTTTTTCAAGAAAGAC 220122
|||||

Search completed: September 26, 2005, 05:02:37
Job time : 364.342 secs

No.	Score	Match	Length	ID	Description
1	411	100.0	411	6	AX040184
2	388.2	94.5	6356	14	AX040184 Sequence
3	373.8	90.9	6355	6	AB078435
4	373.8	90.9	6355	6	AX040177 Sequence
5	373.8	90.9	6355	6	AX795380
6	364.8	88.8	6355	14	TMGCG
7	364.8	88.8	776	14	TMI1229095
8	360.2	87.6	876	14	TMI1229096
9	359.4	87.4	1413	14	TMI1229097
10	336.2	81.8	7685	6	TMGMPCEA
11	336.2	81.8	7685	6	BD263877
12	336.2	81.8	7685	6	AR370290 Sequence
13	336.2	81.8	7685	6	AR370293 Sequence
14	336.2	81.8	7685	6	AR370317 Sequence
15	336.2	81.8	7685	6	AR370320 Sequence
16	336.2	81.8	7685	6	AR435697 Sequence
17	336.2	81.8	7685	6	AX045755 Sequence
18	336.2	81.8	7686	6	BD263878
19	336.2	81.8	7686	6	AR370291 Sequence
20	336.2	81.8	7686	6	AR370294 Sequence

Qy 183 AAGTGTACGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAAATC 242
Db 5622 AAGTGTACGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAAATC 5681

Qy 243 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCTGTGACGCTGATC 302
Db 5682 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCTGTGACGCTGATC 5741

Qy 303 AATCTGTGTACAAATGCAATTAAGTAAACCAAGTTTCAAAACGCAACCAAGCTAGGACACAGTC 362
Db 5742 AATCTGTGTACAAATGCAATTAAGTAAACCAAGTTTCAAAACGCAACCAAGCTAGGACACAGTC 5801

Qy 363 CAACAGCAATTTGCGGATGCTGCAACCTGTGCTAGTATGACAGTGA 411
Db 5802 CAACAGCAATTTGCGGATGCTGCAACCTGTGCTAGTATGACAGTGA 5850

RESULT 3
AX040177 LOCUS AX040177 6355 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 5 from Patent WO0063397.
ACCESSION AX040177
VERSION AX040177.1 GI:11230127
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Meulwaeter, F., Cornelisse, Jacobs, J., van Eldik, G. and Metzlaaff, M.
TITLE Methods and means for delivering inhibitory rna to plants and applications thereof
JOURNAL Patent: WO 0063397-A 5 26-OCT-2000;
Aventis CropScience N.V. (BE)
FEATURES Location/Qualifiers
source 1. 6355
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="cDNA copy of the nucleotide sequence of the genome of TMV-U2"

ORIGIN
Query Match 90.9%; Score 373.8; DB 6; Length 6355;
Best Local Similarity 94.6%; Pred. No. 4.1e-75;
Matches 387; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 3 CTCGCCAATTGAACCTCACTGAAAGTTGTTGATGAGTTCGTAGATGAGTACCGATGGC 62
Db 5441 CTCGCCAATTGAACCTCACTGAAAGTTGTTGATGAGTTCGTAGATGAGTACCGAATGGC 5500

Qy 63 TGTGAACTCGAAAGTTCCGGAAACAAAGAGAGTGTAGTAAATGTTAAATGTTAAATA 122
Db 5501 TGTGAACTCGAAAGTTCCGGAAACAAAGAGAGTGTAGTAAATGTTAAATGTTAAATA 5560

Qy 123 TAAGAAATAAATAATAGTGTAAAGAGGTTTGAAGTTTGAAGAAATTTGAGGAAATATGT 182
Db 5561 TAAGAAATAAATAATAGTGTAAAGAGGTTTGAAGTTTGAAGAAATTTGAGGAAATATGT 5620

Qy 183 AAGTGTACGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAAATC 242
Db 5621 AAGTGTACGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAAATC 5680

Qy 243 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCTGTGACGCTGATC 302
Db 5681 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCTGTGACGCTGATC 5740

Qy 303 AAGTGTACGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAAATC 242
Db 5621 AAGTGTACGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAAATC 5680

Qy 243 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCTGTGACGCTGATC 302
Db 5681 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCTGTGACGCTGATC 5740

Qy 303 AATCTGTGTACAAATGCAATTAAGTAAACCAAGTTTCAAAACGCAACCAAGCTAGGACACAGTC 362
Db 5741 AATCTGTGTACAAATGCAATTAAGTAAACCAAGTTTCAAAACGCAACCAAGCTAGGACACAGTC 5800

Qy 363 CAACAGCAATTTGCGGATGCTGCAACCTGTGCTAGTATGACAGTGA 411
Db 5801 CAACAGCAATTTGCGGATGCTGCAACCTGTGCTAGTATGACAGTGA 5849

RESULT 4
AX795380 LOCUS AX795380 6355 bp mRNA linear PAT 04-OCT-2003
DEFINITION Sequence 7 from Patent WO03052108.
ACCESSION AX795380
VERSION AX795380.1 GI:37516053
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Metzlaaff, M.H., Gosselle, V.M., Meulwaeter, F. and Fache, I.C.
TITLE Improved methods and means for delivering inhibitory rna to plants and applications thereof
JOURNAL Patent: WO 03052108-A 7 26-JUN-2003;
Bayer BioScience N.V. (BE)
FEATURES Location/Qualifiers
source 1. 6355
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="cDNA sequence of the genome of TMV-U2"

ORIGIN
Query Match 90.9%; Score 373.8; DB 6; Length 6355;
Best Local Similarity 94.6%; Pred. No. 4.1e-75;
Matches 387; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 3 CTCGCCAATTGAACCTCACTGAAAGTTGTTGATGAGTTCGTAGATGAGTACCGATGGC 62
Db 5441 CTCGCCAATTGAACCTCACTGAAAGTTGTTGATGAGTTCGTAGATGAGTACCAATGGC 5500

Qy 63 TGTGAACTCGAAAGTTCCGGAAACAAAGAGAGTGTAGTAAATGTTAAATGTTAAATA 122
Db 5501 TGTGAACTCGAAAGTTCCGGAAACAAAGAGAGTGTAGTAAATGTTAAATGTTAAATA 5560

Qy 123 TAAGAAATAAATAATAGTGTAAAGAGGTTTGAAGTTTGAAGAAATTTGAGGAAATATGT 182
Db 5561 TAAGAAATAAATAATAGTGTAAAGAGGTTTGAAGTTTGAAGAAATTTGAGGAAATATGT 5620

Qy 183 AAGTGTACGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAAATC 242
Db 5621 AAGTGTACGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAAATC 5680

Qy 243 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCTGTGACGCTGATC 302
Db 5681 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCTGTGACGCTGATC 5740

Qy 303 AATCTGTGTACAAATGCAATTAAGTAAACCAAGTTTCAAAACGCAACCAAGCTAGGACACAGTC 362
Db 5741 AATCTGTGTACAAATGCAATTAAGTAAACCAAGTTTCAAAACGCAACCAAGCTAGGACACAGTC 5800

Qy 363 CAACAGCAATTTGCGGATGCTGCAACCTGTGCTAGTATGACAGTGA 411
Db 5801 CAACAGCAATTTGCGGATGCTGCAACCTGTGCTAGTATGACAGTGA 5849

RESULT 5
TMCGG LOCUS TMCGG 6355 bp ss-RNA linear VRL 03-AUG-1993
DEFINITION Tobacco mild green mosaic virus complete genome.
ACCESSION M34077 M22483
VERSION M34077.1 GI:335243
KEYWORDS coat protein.
ORGANISM Tobacco mild green mosaic virus
SOURCE Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE 1 (bases 6127 to 6355)
AUTHORS Garcia-Arenal, F.
TITLE Sequence and structure at the genome 3' end of the U2-strain of tobacco mosaic virus, a histidine-accepting tobamovirus

JOURNAL MEDLINE PUBMED	Virology 167 (1), 201-206 (1988)	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL
REFERENCE	89045644 3188396	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST
AUTHORS	2 (bases 1 to 6355)	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV
TITLE	Solis,I. and Garcia-Arenal,F.	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY
COMMENT	The complete nucleotide sequence of the genomic RNA of the tobamovirus tobacco mild green mosaic virus	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD
JOURNAL MEDLINE PUBMED	Virology 177 (2), 553-558 (1990)	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG
FEATURES	90320127 2371769	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL
source	Original source text: Tobacco mild green mosaic virus (strain U2-TMV), cDNA to viral RNA, from N.tabacum cv. Samsum.	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL
	Draft entry and computer-readable sequence for [1] kindly submitted by F.Garcia-Arenal, 08-MAY-1990, for release after publication.	QSKYEGAVTTTISVQSRSSSEMIGKGVLSVSKPLKGIKIVTFOADKFELEKGYKN
	Location/Qualifiers	VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYITVVDLPVQ
	1. .6355	IISDLSLSFLLEMYVEAGSR"
	/organism="Tobacco mild green mosaic virus"	<71. .3403
	/mol_type="genomic RNA"	exon
	/db_xref="taxon:12241"	exon
	join(71. .3403,3407. .4900)	CDS
	/note="183 kDa protein"	
	/codon_start=1	
	/protein_id="AAA47934.1"	
	/db_xref="GI:335244"	
	/translation="MAHISQSIISNALLSVSGKNTLVNDLARRMYDTAVEEFNARDR	
	RPKNFNSKTISEEOTLIVSNAYPEFOITFYNTQNAVSLAGGLRALSLEYLMLQVPYG	
	SPTYDIGNFAAEKGRDYYVCHCCPNLDIRIMRHEGQKDSIEMYLRSLSRKNKVP	
	EFQREANRVAEAEVNECCSKTFQDCRIHPENSGRYAVALHSLYDIPVHEFGAALI	
	SKNHVCAASIALLEADQTEVTLNEIGATFKREGDDVSPFPADRSITLNSHYKN	
	ILHVVKSPASSRIYVFEFLVTRVNTWCKETKYDVTYILKSPQVSCDSQDFE	
	AMEDAFYKTLAMFENTERALFROTASVNFVFPKMKDMVIVPLFEGSITSKMTRSEV	
	IVNRDFTVTLNHRITQOALVNGKFCRLDKTTSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LSMTFFQTLKLAALQDQIVMGRKFLCDKTLSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL	
	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST	
	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV	
	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY	
	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD	
	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG	
	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL	
	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL	
	QSKYEGAVTTTISVQSRSSSEMIGKGVLSVSKPLKGIKIVTFOADKFELEKGYKN	
	VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYITVVDLPVQ	
	IISDLSLSFLLEMYVEAGSRQIQMDAVFKGNLVAITPKSGDFPDLPQFYDCLP	
	GNSTILNKYDAVTRMNRDNLNVDKDFSKSIPMPKVKPCLEPVLRTAAEPRAA	
	GLLENVAMIKRNFNAPDLTGTIDIESASVVDKFDSPYIKKEKTKNIAGVMTKD	
	SMWRLENRREVLDDLANYNFTDLPADIDQVKHMIKAQPKQLDLSIQNEYPALQTV	
	YHSKQINGILAGSELTLLLEAFDSKKFLPFTFKTPQIQERPSDLSHVPMDVLEL	
	DISKYDSQNEPHECAVEYEWKRLGNEFLAEVWKQGRKTTLDYIAGIKTCLWYR	
	KSGDVTTFIGNITIIAACLSGLMPKWKVIGKAFCDGSDVLPFKGLDFPDQSCANLM	
	WNFRKDYRKRYGFCFYQIHHDKGAIYVYDPLKLSKLGAKHIDYDLEELRVS	
	CDVACSLGNWALKEVHTKAIDGSAFNCVKNKFLCDKFLFRTLFLNGC"	
	71. .3406	
	/note="126 kDa protein"	
	/codon_start=1	
	/protein_id="AAA47935.1"	
	/db_xref="GI:335245"	
	/translation="MAHISQSIISNALLSVSGKNTLVNDLARRMYDTAVEEFNARDR	
	RPKNFNSKTISEEOTLIVSNAYPEFOITFYNTQNAVSLAGGLRALSLEYLMLQVPYG	
	SPTYDIGNFAAEKGRDYYVCHCCPNLDIRIMRHEGQKDSIEMYLRSLSRKNKVP	
	EFQREANRVAEAEVNECCSKTFQDCRIHPENSGRYAVALHSLYDIPVHEFGAALI	
	SKNHVCAASIALLEADQTEVTLNEIGATFKREGDDVSPFPADRSITLNSHYKN	
	ILHVVKSPASSRIYVFEFLVTRVNTWCKETKYDVTYILKSPQVSCDSQDFE	
	AMEDAFYKTLAMFENTERALFROTASVNFVFPKMKDMVIVPLFEGSITSKMTRSEV	
	IVNRDFTVTLNHRITQOALVNGKFCRLDKTTSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LSMTFFQTLKLAALQDQIVMGRKFLCDKTLSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL	
	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST	
	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV	
	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY	
	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD	
	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG	
	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL	
	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL	
	QSKYEGAVTTTISVQSRSSSEMIGKGVLSVSKPLKGIKIVTFOADKFELEKGYKN	
	VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYITVVDLPVQ	
	IISDLSLSFLLEMYVEAGSRQIQMDAVFKGNLVAITPKSGDFPDLPQFYDCLP	
	GNSTILNKYDAVTRMNRDNLNVDKDFSKSIPMPKVKPCLEPVLRTAAEPRAA	
	GLLENVAMIKRNFNAPDLTGTIDIESASVVDKFDSPYIKKEKTKNIAGVMTKD	
	SMWRLENRREVLDDLANYNFTDLPADIDQVKHMIKAQPKQLDLSIQNEYPALQTV	
	YHSKQINGILAGSELTLLLEAFDSKKFLPFTFKTPQIQERPSDLSHVPMDVLEL	
	DISKYDSQNEPHECAVEYEWKRLGNEFLAEVWKQGRKTTLDYIAGIKTCLWYR	
	KSGDVTTFIGNITIIAACLSGLMPKWKVIGKAFCDGSDVLPFKGLDFPDQSCANLM	
	WNFRKDYRKRYGFCFYQIHHDKGAIYVYDPLKLSKLGAKHIDYDLEELRVS	
	CDVACSLGNWALKEVHTKAIDGSAFNCVKNKFLCDKFLFRTLFLNGC"	
	71. .3406	
	/note="126 kDa protein"	
	/codon_start=1	
	/protein_id="AAA47935.1"	
	/db_xref="GI:335245"	
	/translation="MAHISQSIISNALLSVSGKNTLVNDLARRMYDTAVEEFNARDR	
	RPKNFNSKTISEEOTLIVSNAYPEFOITFYNTQNAVSLAGGLRALSLEYLMLQVPYG	
	SPTYDIGNFAAEKGRDYYVCHCCPNLDIRIMRHEGQKDSIEMYLRSLSRKNKVP	
	EFQREANRVAEAEVNECCSKTFQDCRIHPENSGRYAVALHSLYDIPVHEFGAALI	
	SKNHVCAASIALLEADQTEVTLNEIGATFKREGDDVSPFPADRSITLNSHYKN	
	ILHVVKSPASSRIYVFEFLVTRVNTWCKETKYDVTYILKSPQVSCDSQDFE	
	AMEDAFYKTLAMFENTERALFROTASVNFVFPKMKDMVIVPLFEGSITSKMTRSEV	
	IVNRDFTVTLNHRITQOALVNGKFCRLDKTTSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LSMTFFQTLKLAALQDQIVMGRKFLCDKTLSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL	
	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST	
	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV	
	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY	
	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD	
	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG	
	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL	
	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL	
	QSKYEGAVTTTISVQSRSSSEMIGKGVLSVSKPLKGIKIVTFOADKFELEKGYKN	
	VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYITVVDLPVQ	
	IISDLSLSFLLEMYVEAGSRQIQMDAVFKGNLVAITPKSGDFPDLPQFYDCLP	
	GNSTILNKYDAVTRMNRDNLNVDKDFSKSIPMPKVKPCLEPVLRTAAEPRAA	
	GLLENVAMIKRNFNAPDLTGTIDIESASVVDKFDSPYIKKEKTKNIAGVMTKD	
	SMWRLENRREVLDDLANYNFTDLPADIDQVKHMIKAQPKQLDLSIQNEYPALQTV	
	YHSKQINGILAGSELTLLLEAFDSKKFLPFTFKTPQIQERPSDLSHVPMDVLEL	
	DISKYDSQNEPHECAVEYEWKRLGNEFLAEVWKQGRKTTLDYIAGIKTCLWYR	
	KSGDVTTFIGNITIIAACLSGLMPKWKVIGKAFCDGSDVLPFKGLDFPDQSCANLM	
	WNFRKDYRKRYGFCFYQIHHDKGAIYVYDPLKLSKLGAKHIDYDLEELRVS	
	CDVACSLGNWALKEVHTKAIDGSAFNCVKNKFLCDKFLFRTLFLNGC"	
	71. .3406	
	/note="126 kDa protein"	
	/codon_start=1	
	/protein_id="AAA47935.1"	
	/db_xref="GI:335245"	
	/translation="MAHISQSIISNALLSVSGKNTLVNDLARRMYDTAVEEFNARDR	
	RPKNFNSKTISEEOTLIVSNAYPEFOITFYNTQNAVSLAGGLRALSLEYLMLQVPYG	
	SPTYDIGNFAAEKGRDYYVCHCCPNLDIRIMRHEGQKDSIEMYLRSLSRKNKVP	
	EFQREANRVAEAEVNECCSKTFQDCRIHPENSGRYAVALHSLYDIPVHEFGAALI	
	SKNHVCAASIALLEADQTEVTLNEIGATFKREGDDVSPFPADRSITLNSHYKN	
	ILHVVKSPASSRIYVFEFLVTRVNTWCKETKYDVTYILKSPQVSCDSQDFE	
	AMEDAFYKTLAMFENTERALFROTASVNFVFPKMKDMVIVPLFEGSITSKMTRSEV	
	IVNRDFTVTLNHRITQOALVNGKFCRLDKTTSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LSMTFFQTLKLAALQDQIVMGRKFLCDKTLSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL	
	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST	
	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV	
	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY	
	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD	
	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG	
	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL	
	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL	
	QSKYEGAVTTTISVQSRSSSEMIGKGVLSVSKPLKGIKIVTFOADKFELEKGYKN	
	VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYITVVDLPVQ	
	IISDLSLSFLLEMYVEAGSRQIQMDAVFKGNLVAITPKSGDFPDLPQFYDCLP	
	GNSTILNKYDAVTRMNRDNLNVDKDFSKSIPMPKVKPCLEPVLRTAAEPRAA	
	GLLENVAMIKRNFNAPDLTGTIDIESASVVDKFDSPYIKKEKTKNIAGVMTKD	
	SMWRLENRREVLDDLANYNFTDLPADIDQVKHMIKAQPKQLDLSIQNEYPALQTV	
	YHSKQINGILAGSELTLLLEAFDSKKFLPFTFKTPQIQERPSDLSHVPMDVLEL	
	DISKYDSQNEPHECAVEYEWKRLGNEFLAEVWKQGRKTTLDYIAGIKTCLWYR	
	KSGDVTTFIGNITIIAACLSGLMPKWKVIGKAFCDGSDVLPFKGLDFPDQSCANLM	
	WNFRKDYRKRYGFCFYQIHHDKGAIYVYDPLKLSKLGAKHIDYDLEELRVS	
	CDVACSLGNWALKEVHTKAIDGSAFNCVKNKFLCDKFLFRTLFLNGC"	
	71. .3406	
	/note="126 kDa protein"	
	/codon_start=1	
	/protein_id="AAA47935.1"	
	/db_xref="GI:335245"	
	/translation="MAHISQSIISNALLSVSGKNTLVNDLARRMYDTAVEEFNARDR	
	RPKNFNSKTISEEOTLIVSNAYPEFOITFYNTQNAVSLAGGLRALSLEYLMLQVPYG	
	SPTYDIGNFAAEKGRDYYVCHCCPNLDIRIMRHEGQKDSIEMYLRSLSRKNKVP	
	EFQREANRVAEAEVNECCSKTFQDCRIHPENSGRYAVALHSLYDIPVHEFGAALI	
	SKNHVCAASIALLEADQTEVTLNEIGATFKREGDDVSPFPADRSITLNSHYKN	
	ILHVVKSPASSRIYVFEFLVTRVNTWCKETKYDVTYILKSPQVSCDSQDFE	
	AMEDAFYKTLAMFENTERALFROTASVNFVFPKMKDMVIVPLFEGSITSKMTRSEV	
	IVNRDFTVTLNHRITQOALVNGKFCRLDKTTSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LSMTFFQTLKLAALQDQIVMGRKFLCDKTLSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL	
	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST	
	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV	
	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY	
	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD	
	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG	
	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL	
	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL	
	QSKYEGAVTTTISVQSRSSSEMIGKGVLSVSKPLKGIKIVTFOADKFELEKGYKN	
	VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYITVVDLPVQ	
	IISDLSLSFLLEMYVEAGSRQIQMDAVFKGNLVAITPKSGDFPDLPQFYDCLP	
	GNSTILNKYDAVTRMNRDNLNVDKDFSKSIPMPKVKPCLEPVLRTAAEPRAA	
	GLLENVAMIKRNFNAPDLTGTIDIESASVVDKFDSPYIKKEKTKNIAGVMTKD	
	SMWRLENRREVLDDLANYNFTDLPADIDQVKHMIKAQPKQLDLSIQNEYPALQTV	
	YHSKQINGILAGSELTLLLEAFDSKKFLPFTFKTPQIQERPSDLSHVPMDVLEL	
	DISKYDSQNEPHECAVEYEWKRLGNEFLAEVWKQGRKTTLDYIAGIKTCLWYR	
	KSGDVTTFIGNITIIAACLSGLMPKWKVIGKAFCDGSDVLPFKGLDFPDQSCANLM	
	WNFRKDYRKRYGFCFYQIHHDKGAIYVYDPLKLSKLGAKHIDYDLEELRVS	
	CDVACSLGNWALKEVHTKAIDGSAFNCVKNKFLCDKFLFRTLFLNGC"	
	71. .3406	
	/note="126 kDa protein"	
	/codon_start=1	
	/protein_id="AAA47935.1"	
	/db_xref="GI:335245"	
	/translation="MAHISQSIISNALLSVSGKNTLVNDLARRMYDTAVEEFNARDR	
	RPKNFNSKTISEEOTLIVSNAYPEFOITFYNTQNAVSLAGGLRALSLEYLMLQVPYG	
	SPTYDIGNFAAEKGRDYYVCHCCPNLDIRIMRHEGQKDSIEMYLRSLSRKNKVP	
	EFQREANRVAEAEVNECCSKTFQDCRIHPENSGRYAVALHSLYDIPVHEFGAALI	
	SKNHVCAASIALLEADQTEVTLNEIGATFKREGDDVSPFPADRSITLNSHYKN	
	ILHVVKSPASSRIYVFEFLVTRVNTWCKETKYDVTYILKSPQVSCDSQDFE	
	AMEDAFYKTLAMFENTERALFROTASVNFVFPKMKDMVIVPLFEGSITSKMTRSEV	
	IVNRDFTVTLNHRITQOALVNGKFCRLDKTTSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LSMTFFQTLKLAALQDQIVMGRKFLCDKTLSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL	
	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST	
	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV	
	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY	
	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD	
	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG	
	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL	
	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL	
	QSKYEGAVTTTISVQSRSSSEMIGKGVLSVSKPLKGIKIVTFOADKFELEKGYKN	
	VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYITVVDLPVQ	
	IISDLSLSFLLEMYVEAGSRQIQMDAVFKGNLVAITPKSGDFPDLPQFYDCLP	
	GNSTILNKYDAVTRMNRDNLNVDKDFSKSIPMPKVKPCLEPVLRTAAEPRAA	
	GLLENVAMIKRNFNAPDLTGTIDIESASVVDKFDSPYIKKEKTKNIAGVMTKD	
	SMWRLENRREVLDDLANYNFTDLPADIDQVKHMIKAQPKQLDLSIQNEYPALQTV	
	YHSKQINGILAGSELTLLLEAFDSKKFLPFTFKTPQIQERPSDLSHVPMDVLEL	
	DISKYDSQNEPHECAVEYEWKRLGNEFLAEVWKQGRKTTLDYIAGIKTCLWYR	
	KSGDVTTFIGNITIIAACLSGLMPKWKVIGKAFCDGSDVLPFKGLDFPDQSCANLM	
	WNFRKDYRKRYGFCFYQIHHDKGAIYVYDPLKLSKLGAKHIDYDLEELRVS	
	CDVACSLGNWALKEVHTKAIDGSAFNCVKNKFLCDKFLFRTLFLNGC"	
	71. .3406	
	/note="126 kDa protein"	
	/codon_start=1	
	/protein_id="AAA47935.1"	
	/db_xref="GI:335245"	
	/translation="MAHISQSIISNALLSVSGKNTLVNDLARRMYDTAVEEFNARDR	
	RPKNFNSKTISEEOTLIVSNAYPEFOITFYNTQNAVSLAGGLRALSLEYLMLQVPYG	
	SPTYDIGNFAAEKGRDYYVCHCCPNLDIRIMRHEGQKDSIEMYLRSLSRKNKVP	
	EFQREANRVAEAEVNECCSKTFQDCRIHPENSGRYAVALHSLYDIPVHEFGAALI	
	SKNHVCAASIALLEADQTEVTLNEIGATFKREGDDVSPFPADRSITLNSHYKN	
	ILHVVKSPASSRIYVFEFLVTRVNTWCKETKYDVTYILKSPQVSCDSQDFE	
	AMEDAFYKTLAMFENTERALFROTASVNFVFPKMKDMVIVPLFEGSITSKMTRSEV	
	IVNRDFTVTLNHRITQOALVNGKFCRLDKTTSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LSMTFFQTLKLAALQDQIVMGRKFLCDKTLSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL	
	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST	
	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV	
	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY	
	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD	
	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG	
	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL	
	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL	
	QSKYEGAVTTTISVQSRSSSEMIGKGVLSVSKPLKGIKIVTFOADKFELEKGYKN	
	VNTVHEIQGETFEDVSLVRLTATPLTLISKSSPHVLVALTRHTKSFKYITVVDLPVQ	
	IISDLSLSFLLEMYVEAGSRQIQMDAVFKGNLVAITPKSGDFPDLPQFYDCLP	
	GNSTILNKYDAVTRMNRDNLNVDKDFSKSIPMPKVKPCLEPVLRTAAEPRAA	
	GLLENVAMIKRNFNAPDLTGTIDIESASVVDKFDSPYIKKEKTKNIAGVMTKD	
	SMWRLENRREVLDDLANYNFTDLPADIDQVKHMIKAQPKQLDLSIQNEYPALQTV	
	YHSKQINGILAGSELTLLLEAFDSKKFLPFTFKTPQIQERPSDLSHVPMDVLEL	
	DISKYDSQNEPHECAVEYEWKRLGNEFLAEVWKQGRKTTLDYIAGIKTCLWYR	
	KSGDVTTFIGNITIIAACLSGLMPKWKVIGKAFCDGSDVLPFKGLDFPDQSCANLM	
	WNFRKDYRKRYGFCFYQIHHDKGAIYVYDPLKLSKLGAKHIDYDLEELRVS	
	CDVACSLGNWALKEVHTKAIDGSAFNCVKNKFLCDKFLFRTLFLNGC"	
	71. .3406	
	/note="126 kDa protein"	
	/codon_start=1	
	/protein_id="AAA47935.1"	
	/db_xref="GI:335245"	
	/translation="MAHISQSIISNALLSVSGKNTLVNDLARRMYDTAVEEFNARDR	
	RPKNFNSKTISEEOTLIVSNAYPEFOITFYNTQNAVSLAGGLRALSLEYLMLQVPYG	
	SPTYDIGNFAAEKGRDYYVCHCCPNLDIRIMRHEGQKDSIEMYLRSLSRKNKVP	
	EFQREANRVAEAEVNECCSKTFQDCRIHPENSGRYAVALHSLYDIPVHEFGAALI	
	SKNHVCAASIALLEADQTEVTLNEIGATFKREGDDVSPFPADRSITLNSHYKN	
	ILHVVKSPASSRIYVFEFLVTRVNTWCKETKYDVTYILKSPQVSCDSQDFE	
	AMEDAFYKTLAMFENTERALFROTASVNFVFPKMKDMVIVPLFEGSITSKMTRSEV	
	IVNRDFTVTLNHRITQOALVNGKFCRLDKTTSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LSMTFFQTLKLAALQDQIVMGRKFLCDKTLSELIVSEVGRKFGVNTARSEWDVRAILOP	
	LDVSENALKIKIPDLVYVTKDRFVAEYTKSEELPHLIDIKKDLBEAEQMYDALSLSIL	
	KGADNFDIAFKDMCKALDVS PDVAARVIVAAENRSGLTITLTDKPTPEENVAKALKST	
	ASEAVVCLEPTSEEVNKNFKSIAEGRPLVCAESHGGLTNANLEHQSLESNDFHKACV	
	DSVITKQMASVVTYSLKQVQMKNYVDSLAAISATVSNLCKSLKSDVGVDSDSREKY	
	GWMDVTLKKWLLPKPAAGHSGWVLDYKGRMFTALLSYEGDRMVTESDWRRAVVSDD	
	MVSDIAKQNLKRTMDRGPHEPTAKMVLVDGVPGCCYKGFDPERDLDGLTVPFG	
	KQAAAMRRRANSSGLIRATMDNVRTVDSLMLHPKPRSHKRLFDI EGLMLHTGCVNFL	
	VLISGCDIAYIGDTQOIIPFINRVQNPFPYPKPEKLOQDVEVEMERTTLRCPCGVNFFL	
	QSKYEGAVTTTISVQSRSSSEMIGKGVLS	


```
ACCESSION AJ429095
VERSION AJ429095.1 GI:28170589
KEYWORDS capsid protein; cp gene.
SOURCE Tobacco mild green mosaic virus
ORGANISM Tobacco mild green mosaic virus
REFERENCE 1 Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
AUTHORS Letschert,B., Adam,G., Lesemann,D., Willingmann,P. and Heinze,C.
TITLE Detection and differentiation of serologically cross-reacting
Tobamoviruses of economical importance by RT-PCR and RT-PCR-RFLP
JOURNAL J. Virol. Methods 106 (1), 1-10 (2002)
MEDLINE 22255451
PUBMED 12367724
REFERENCE 2 (bases 1 to 776)
AUTHORS Willingmann,P.
TITLE Direct Submision
JOURNAL Submitted (28-JAN-2002) Willingmann P., Abteilung Pflanzenschutz,
Institut fuer Angewandte Botanik, Ohnhorstr.18, 22609 Hamburg,
GERMANY
FEATURES             Location/Qualifiers
     source           1. .776
                     /organism="Tobacco mild green mosaic virus"
                     /viroion
                     /mol_type="genomic RNA"
                     /isolate="DSMZ PV-110"
                     /db_xref="taxon:12241"
                     /country="France:Corsica"
     gene             201. .680
                     /gene="cp"
     CDS               201. .680
                     /gene="cp"
                     /codon_start=1
                     /product="capsid protein"
                     /protein_id="CAD22092.1"
                     /db_xref="GI:28170590"
                     /db_xref="GOA:Q80P17"
                     /db_xref="UniProt/TREMBL:Q80P17"
                     /translation="MPYITNSPQFVYLSSAYADPVQLINLRTNALGNQFOTQOARTT
                     VQOQFADAKVPMSVTRFPASDFVYRNSYDPLITALINSFDRNRRIIEVDNQA
                     PNTEIVNTQVRDDATVAIRASINNLANELVRGTGMFNQAGFETASGLVWTTTPAT"

ORIGIN
Query Match      88.8%; Score 364.8; DB 14; Length 776;
Best Local Similarity 96.9%; Pred. No. 6.1e-73;
Matches 372; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 28 GTTGTGATGATTCGTAGATGAAGTACCGATGGCTGTGAACCTCGAAAGGTTCCGGAAA 87
DB 1 GTTGTGATGATTCCTTCGATGAAGTACCAATGGCTGTGAACCTCGAAAGGTTCCGGAAA 60

QY 88 ACAAAAAGAGAGTGGTAGGTAATAATGTTAATAATAGAAAATAAATAAGTGGTAAG 147
DB 61 ACAAAAAGAAAGTGGTAGGTAATCATGTTAATAATAGAAAATAAATAAGTGGTAAG 120

QY 148 AAGGTTTGAAGTTGAGGAAATTTGAGGATAATGTAAGTGATGACGAGTCTATCGCGTCA 207
DB 121 AAGGTTTGAAGTTGAGGAAATTTGAGGATAATGTAAGTGATGACGAGTCTATCGCGTCA 180

QY 208 TCGAGTAGCTTTAATCAATATGCCTTATACAATCAACTCTCCGAGCCAAATTTGTTTACT 267
DB 181 TCGAGTAGCTTTAATCAATATGCCTTATACAATCAACTCTCCGAGCCAAATTTGTTTACT 240

QY 268 TAAGTTCCGCTTATGCAGATCCTGTGCAGCTGATCAATCTGTGTACAAATGATAGGTA 327
DB 241 TAAGTTCCGCTTATGCAGATCCTGTGCAGCTGATCAATCTGTGTACAAATGATAGGTA 300
```

```

Qy 328 ACCAGTTTCAAACCAACAGCTAGGACAAACAGTCCCAACAGCAATTTGCGGATCCCTGGA 387
Db 301 ACCAGTTTCAAACCAACAGCTAGGACAAACAGTCCCAACAGCAATTTGCGGATCCCTGGA 360
Qy 388 AACCTGTGCTAGTATGACAGTGA 411
Db 361 AACCTGTGCTAGTATGACAGTGA 384

RESULT 8
TM1429097
LOCUS Tobacco mild green mosaic virus cp gene for capsid protein, genomic
RNA, isolate DSMZ PV-120.
ACCESSION AJ429097.1 GI:28170593
VERSION capsid protein; cp gene.
KEYWORDS Tobacco mild green mosaic virus
SOURCE Tobacco mild green mosaic virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE 1
AUTHORS Letschert,B., Adam,G., Lesemann,D., Willingmann,P. and Heinze,C.
TITLE Detection and differentiation of serologically cross-reacting
Tobamoviruses of economical importance by RT-PCR and RT-PCR-RFLP
J. Virol. Methods 106 (1), 1-10 (2002)
2255451
MEDLINE 12367724
PUBMED 2
(bases 1 to 876)
Willingmann,P.
AUTHORS Direct Submission
TITLE Submitted (28-JAN-2002) Willingmann P., Abteilung Pflanzenschutz,
JOURNAL Institut fuer Angewandte Botanik, Ohnhorststr.18, 22609 Hamburg,
GERMANY
FEATURES
source Location/Qualifiers
1. .876
/organism="Tobacco mild green mosaic virus"
/virus
/mol_type="genomic RNA"
/isolate="DSMZ PV-120"
/db_xref="taxon:12241"
/country="Italy"
202. .681
/gene="cp"
202. .681
/gene="cp"
/codon_start=1
/product="capsid protein"
/protein_id="CAD22094.1"
/db_xref="GI:28170594"
/db_xref="GOA:Q80P15"
/db_xref="UniProt/TREMBL:Q80P15"
/translation="MPYITNSPKFVYLSSAYADPVOLINLTNALGNQFQTOQARTT
VQQQFADAWKPVPSITVRFPSADPVYRYNSTLDPLITALLNSPDTNRRIIEVDNQA
PNTTEIVNATQRVDDATVAIRASINNLANELVRGTGMFNQAGPFTASGLVWTTTSPAT"

ORIGIN
Query Match 87.6%; Score 360.2; DB 14; Length 876;
Best Local Similarity 97.7%; Pred. No. 6.7e-72;
Matches 376; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 28 GTTCGTGATGATTCGTAGTAGTAAAGTACCGATGGC-TGTGAACTCCGAAAGGTTCCGGAA 86
Db 1 GTTCGTGAGGATTCGTGATGAAAGTACCGATGGCGTGTGAARTCCGAAAGGTTCCGGAA 60

Qy 87 AACAAAAAGAGGTGGTAGTAAATGTTTAAATAAAGAAAAATAAATAAGTGGTAA 146
Db 61 AACAAAAAGAGGTGGTAGTAAATGTTTAAATAAAGAAAAATAAATAAGTGGTAA 120

Qy 147 GAAGGGTTTGAAGTTGAGGAAATGAGGATAATGTAGTGATGACCGAGTCTATCGGTC 206
Db 121 GAAGGGTTTGAAGTTGAGGAAATGAGGATAATGTAGTGATGACCGAGTCTATCGGTC 180

Qy 207 ATCAGGTACGTTTAAATCAATATGCTTATCAATCAACTCTCCGAGGCAATTTGTTTAC 266

```

```

Db 181 ATCGAGTACGTTTAAATCAATATGCTTATCAATCAACTCTCCGAGCAATTTGTTTAC 240
Qy 267 TTAAGTTCCTGCTATGAGATCCTGTGACAGTCAATCAATCTGTGTACAAATGCAATAGGT 326
Db 241 TTAAGTTCCTGCTATGAGATCCTGTGACAGTCAATCAATCTGTGTACAAATGCAATGGGT 300
Qy 327 AACAGTTTCAAACGCAACAGCTAGACAAACAGTCCCAACAGCAATTTGCGGATCCCTGG 386
Db 301 AACAGTTTCAAACGCAACAGCTAGACAAACAGTCCCAACAGCAATTTGCGGATCCCTGG 360
Qy 387 AACCTGTGCTAGTATGACAGTGA 411
Db 361 AACCTGTGCTAGTATGACAGTGA 385

RESULT 9
TMGMPCPA
LOCUS Tobacco mild green mosaic virus movement and coat protein genes,
complete cds.
ACCESSION M34236
VERSION M34236.1 GI:335248
KEYWORDS coat protein; movement protein.
SOURCE Tobacco mild green mosaic virus
ORGANISM Tobacco mild green mosaic virus
REFERENCE 1
AUTHORS Nejidat,A., Cellier,P., Holt,C.A., Gafny,R., Eggenberger,A.L. and
Beachy,R.N.
TITLE Transfer of the movement protein gene between two tobamoviruses:
influence on local lesion development
Virology 180 (1), 318-326 (1991)
91082424
PUBMED 1984654
COMMENT Original source text: Tobacco mild green mosaic virus (strain PV
228), cDNA to viral RNA.
Draft entry and computer-readable sequence for [1] kindly submitted
by A.Nejidat, 11-MAY-1990.
FEATURES
source Location/Qualifiers
1. .1413
/organism="Tobacco mild green mosaic virus"
/mol_type="genomic RNA"
/db_xref="taxon:12241"
51. .851
/codon_start=1
/product="movement protein"
/protein_id="AAA47938.1"
/db_xref="GI:335249"
/translation="MAVSLRDTVKISFNLKQDELPAFWTKVKSVRISTVDKIMA
VKDLSDDVLKGLVKNGYVCLAGLVSGEWNLPDNCRGVSVICIVDKMKRSNE
ATLGAHAPACKNFSLIPNYSITSEDAEKNPQVLVNIKGVAMEEGYCLSEFV
SICVVAHKNVYKGLRERILRVTDSDPIELTEKVEEVEFVPMVKLERFRKTKGKK
RKEKKRVVGVGNSVNNKINNKGKGLKVEIEDNVSDDESIASSSTF"
857. 1336
/codon_start=1
/product="coat protein"
/protein_id="AAA47939.1"
/db_xref="GI:335250"
/translation="MPYITNSPKFVYLSSAYADPVOLINLTNALGNQFQTOQARTT
VQQQFADAWKPVPSITVRFPSADPVYRYNSTLDPLITALLNSPDTNRRIIEVDNQA
PNTTEIVNATQRVDDATVAIRASINNLANELVRGTGMFNQAGPFTASGLVWTTTSPAT"

ORIGIN
Query Match 87.4%; Score 359.4; DB 14; Length 1413;
Best Local Similarity 91.8%; Pred. No. 9.6e-72;
Matches 403; Conservative 0; Mismatches 6; Indels 30; Gaps 1;

Qy 3 CTCGCCAATTTGAATCTACTGAAAAAGTTGTTGATGAGTTCGTAGATGAAGTACCGATGGC 62
Db 602 CTCGCCAATTTGAATCTACTGAAAAAGTTGTTGAGGAGTTCGTGAGTGAAGTACCAATGGC 661

Qy 63 TGTGAACTCGAAAGGTTCCGAAAC-----AAA 92

```

```
Db 662 TGTGAAACTCGAAAGGTTCCGGAAACAAAAAAGGGAAAAAAGAAAAA 721
|||
Qy 93 AAAGAGAGTGGTAGCTAATATGTTAATAAAGAAAAATAATATAGTGTGAAGAGGG 152
|||
Db 722 AAAGAGAGTGGTAGCTAATATGTTAATAAAGAAAAATAATATAGTGTGAAGAGGG 781
|||
Qy 153 TTTGAAAGTTGAGGAAATGAGGATAATGTAAGTAGATGACGAGCTATCGCGTCATCGAG 212
|||
Db 782 TTTGAAAGTTGAGGAAATGAGGATAATGTAAGTAGATGACGAGCTATCGCGTCATCGAG 841
|||
Qy 213 TACGTTTTAATCAATATGCTTTATACAACTCACTCCGAGCCAAATTTGTTTACTTAAGT 272
|||
Db 842 TACGTTTTAATCAATATGCTTTATACAACTCACTCCGAGCCAAATTTGTTTACTTAAGT 901
|||
Qy 273 TCCGCTTATGAGATCCTGTGTCAGCTGATCAATCTGTGTACAAATGCAATAGGTAACCG 332
|||
Db 902 TCCGCTTATGAGATCCTGTGTCAGCTGATCAATCTGTGTACAAATGCAATAGGTAACCG 961
|||
Qy 333 TTTCAAAAGCAACAGCTAGGACAAACAGTCCAAAGCAATTTGCGGATGCTCGAAACCT 392
|||
Db 962 TTTCAAAAGCAACAGCTAGGACAAACAGTCCAAAGCAATTTGCGGATGCTCGAAACCT 1021
|||
Qy 393 GTGCTAGTATGACAGTGA 411
|||
Db 1022 GTGCTAGTATGACAGTGA 1040
|||

RESULT 10
BD263877
LOCUS BD263877 7685 bp DNA linear PAT 17-JUL-2003
DEFINITION Viral expression vectors.
ACCESSION BD263877
VERSION BD263877.1 GI:33073645
KEYWORDS JP 2002542828-A/1. (common tobacco)
SOURCE Nicotiana tabacum
ORGANISM Nicotiana tabacum
REFERENCE 1 (bases 1 to 7685)
AUTHORS Fitzmaurice,W.P., Fogue,G.P. and Lindbo,J.A.
TITLES Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
JOURNAL Patent: JP 2002542828-A 1 17-DEC-2002;
LARGE SCALE BIOLOGY CORP
COMMENT OS Nicotiana tabacum (tobacco)
PN JP 2002542828-A/1
PD 17-DEC-2002
PF 04-MAY-2000 JP 2000615766
PR 04-MAY-1999 US 60/132697
PI WAYNE P FITZMAURICE,GREGORY P FOGUE,JOHN A LINDBO PC
C12N15/09,A01H5/00,C12N5/10,C12N15/00,C12N5/00 CC Viral
expression vectors
FH Key Location/Qualifiers
FT source 1..7685
FT /organism='Nicotiana tabacum (tobacco)'.
FEATURES
source Location/Qualifiers
1..7685
/organism='Nicotiana tabacum'
/mol_type='genomic DNA'
/db_xref='taxon:4097'
ORIGIN
Query Match 81.8%; Score 336.2; DB 6; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.5e-66;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAACAAAAAGAGAGTGGTAGTAATAGTTA 118
|||
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACAAAAAGAGAGTGGTAGTAATAGTTA 6740
|||

Qy 119 ATAATAAGAAAAATAATAATAGTGGTAAGAGGGTTTGAAAGTTGAGGAAATGAGGATA 178
|||
Db 6741 ATAATAAGAAAAATAATAATAGTGGTAAGAGGGTTTGAAAGTTGAGGAAATGAGGATA 6800
|||

Qy 179 ATGTAAGTATGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCTTATAC 238
|||
Db 6801 ATGTAAGTATGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCTTATAC 6860
|||

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGACAGATCCTGTGTCAGCT 298
|||
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGACAGATCCTGTGTCAGCT 6920
|||

Qy 299 GATCAATCTGTGTACAAATGCAATGATTAACCAAGTTTCAACAGCAACAGCTAGGACAAAC 358
|||
Db 6921 GATCAATCTGTGTACAAATGCAATGATTAACCAAGTTTCAACAGCAACAGCTAGGACAAAC 6980
|||

Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
|||
Db 6981 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033
|||

RESULT 12
AR370293
LOCUS AR370293 7685 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 25 from patent US 6300133.
```

```
Db 6741 ATAATAAGAAAAATAATAATAGTGGTAAGAGGGTTTGAAAGTTGAGGAAATGAGGATA 6800
|||
Qy 179 ATGTAAGTATGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCTTATAC 238
|||
Db 6801 ATGTAAGTATGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCTTATAC 6860
|||
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGACAGATCCTGTGTCAGCT 298
|||
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGACAGATCCTGTGTCAGCT 6920
|||

Qy 299 GATCAATCTGTGTACAAATGCAATGATTAACCAAGTTTCAACAGCAACAGCTAGGACAAAC 358
|||
Db 6921 GATCAATCTGTGTACAAATGCAATGATTAACCAAGTTTCAACAGCAACAGCTAGGACAAAC 6980
|||

Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
|||
Db 6981 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033
|||

RESULT 11
AR370290
LOCUS AR370290 7685 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 22 from patent US 6300133.
ACCESSION AR370290
VERSION AR370290.1 GI:34606796
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7685)
AUTHORS Lindbo,J.A., Fogue,G.P. and Turpen,T.H.
TITLES RNA transformation vectors derived from an uncapped
single-component RNA virus
JOURNAL Patent: US 6300133-A 22 09-OCT-2001;
FEATURES
source Location/Qualifiers
1..7685
/organism='unknown'
/mol_type='genomic DNA'
ORIGIN
Query Match 81.8%; Score 336.2; DB 6; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.5e-66;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAACAAAAAGAGAGTGGTAGTAATAGTTA 118
|||
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACAAAAAGAGAGTGGTAGTAATAGTTA 6740
|||

Qy 119 ATAATAAGAAAAATAATAATAGTGGTAAGAGGGTTTGAAAGTTGAGGAAATGAGGATA 178
|||
Db 6741 ATAATAAGAAAAATAATAATAGTGGTAAGAGGGTTTGAAAGTTGAGGAAATGAGGATA 6800
|||

Qy 179 ATGTAAGTATGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCTTATAC 238
|||
Db 6801 ATGTAAGTATGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCTTATAC 6860
|||

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGACAGATCCTGTGTCAGCT 298
|||
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGACAGATCCTGTGTCAGCT 6920
|||

Qy 299 GATCAATCTGTGTACAAATGCAATGATTAACCAAGTTTCAACAGCAACAGCTAGGACAAAC 358
|||
Db 6921 GATCAATCTGTGTACAAATGCAATGATTAACCAAGTTTCAACAGCAACAGCTAGGACAAAC 6980
|||

Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
|||
Db 6981 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033
|||

RESULT 12
AR370293
LOCUS AR370293 7685 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 25 from patent US 6300133.
```

ACCESSION AR370293
VERSION AR370293.1 GI:34606799
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7685)
AUTHORS Lindbo, J.A., Pogue, G.P. and Turpen, T.H.
TITLE RNA transformation vectors derived from an uncapped
single-component RNA virus
JOURNAL Patent: US 6300133-A 25 09-OCT-2001;
FEATURES Location/Qualifiers
source 1..7685
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 81.8%; Score 336.2; DB 6; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.5e-66;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAAAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 118
Db 6681 GGCTGTGAAACTCGAAAAAGGTTCCGGAAAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 6740
Qy 119 ATAATAAGAAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTGAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTGAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAAACGAGTTTCAACGCAACAAAGCTAGGACAA 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAAACGAGTTTCAACGCAACAAAGCTAGGACAA 6980
Qy 359 AGTCCACAGCAAAATTCGGGATGCTCGGAAACCTGTGCCTAGTATGACAGTGA 411
Db 6981 AGTCCACAGCAAAATTCGGGATGCTCGGAAACCTGTGCCTAGTATGACAGTGA 7033

RESULT 13
LOCUS AR370317
DEFINITION Sequence 22 from patent US 6300134.
ACCESSION AR370317
VERSION AR370317.1 GI:34606823
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7685)
AUTHORS Lindbo, J.A., Pogue, G.P. and Turpen, T.H.
TITLE RNA transformation vectors derived from a single-component RNA virus and contain an intervening sequence between the cap and the 5' end
JOURNAL Patent: US 6300134-A 22 09-OCT-2001;
FEATURES Location/Qualifiers
source 1..7685
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 81.8%; Score 336.2; DB 6; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.5e-66;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAAAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 118
Db 6681 GGCTGTGAAACTCGAAAAAGGTTCCGGAAAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 6740
Qy 119 ATAATAAGAAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTGAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTGAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAAACGAGTTTCAACGCAACAAAGCTAGGACAA 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAAACGAGTTTCAACGCAACAAAGCTAGGACAA 6980
Qy 359 AGTCCACAGCAAAATTCGGGATGCTCGGAAACCTGTGCCTAGTATGACAGTGA 411
Db 6981 AGTCCACAGCAAAATTCGGGATGCTCGGAAACCTGTGCCTAGTATGACAGTGA 7033

Db 6681 GGCTGTGAAACTCGAAAAAGGTTCCGGAAAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 6740
Qy 119 ATAATAAGAAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTGAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTGAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAAACGAGTTTCAACGCAACAAAGCTAGGACAA 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAAACGAGTTTCAACGCAACAAAGCTAGGACAA 6980
Qy 359 AGTCCACAGCAAAATTCGGGATGCTCGGAAACCTGTGCCTAGTATGACAGTGA 411
Db 6981 AGTCCACAGCAAAATTCGGGATGCTCGGAAACCTGTGCCTAGTATGACAGTGA 7033

RESULT 14
LOCUS AR370320
DEFINITION Sequence 25 from patent US 6300134.
ACCESSION AR370320
VERSION AR370320.1 GI:34606826
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7685)
AUTHORS Lindbo, J.A., Pogue, G.P. and Turpen, T.H.
TITLE RNA transformation vectors derived from a single-component RNA virus and contain an intervening sequence between the cap and the 5' end
JOURNAL Patent: US 6300134-A 25 09-OCT-2001;
FEATURES Location/Qualifiers
source 1..7685
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 81.8%; Score 336.2; DB 6; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.5e-66;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAAAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 118
Db 6681 GGCTGTGAAACTCGAAAAAGGTTCCGGAAAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 6740
Qy 119 ATAATAAGAAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTGAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTGAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAAACGAGTTTCAACGCAACAAAGCTAGGACAA 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAAACGAGTTTCAACGCAACAAAGCTAGGACAA 6980
Qy 359 AGTCCACAGCAAAATTCGGGATGCTCGGAAACCTGTGCCTAGTATGACAGTGA 411
Db 6981 AGTCCACAGCAAAATTCGGGATGCTCGGAAACCTGTGCCTAGTATGACAGTGA 7033

```
RESULT 15
AR435697
LOCUS AR435697 7685 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6656726.
ACCESSION AR435697
VERSION AR435697.1 GI:40198771
KEYWORDS
SOURCE
ORGANISM
Unklassified
REFERENCE 1 (bases 1 to 7685)
AUTHORS Fitzmaurice,W.P., Fogue,G.P. and Lindbo,J.A.
TITLE Viral expression vectors
JOURNAL Patent: US 6656726-A 1 02-DEC-2003;
FEATURES
    source
        1..7685
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Query Match 81.8%; Score 336.2; DB 6; Length 7685;
Best Local Similarity 98.9%; Pred.No.1.5e-66;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAAACAAAAAGAGAGTGGTAGGTAATAATGTGA 118
Db 6681 GGCTGTGAACTCGAAAGGTTCCGGAAACAAAAAGAGAGTGGTAGGTAATAATGTGA 6740
Qy 119 ATAATAAGAAAAATAATAATAGTGGTAAGAGGGTTTGAAAGTTGAGGAAATTCAGGATA 178
Db 6741 ATAATAAGAAAAATAATAATAGTGGTAAGAGGGTTTGAAAGTTGAGGAAATTCAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGTAACCACTTTCAAACGCAACAAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGTAACCACTTTCAAACGCAACAAGCTAGGACAAC 6980
Qy 359 AGTCCACAGCAATTTGCGGATGCTCGAAACCTGTGCCTAGTATGACAGTGA 411
Db 6981 AGTCCACAGCAATTTGCGGATGCTCGAAACCTGTGCCTAGTATGACAGTGA 7033
```

Search completed: September 26, 2005, 03:05:13
Job time : 2890.42 secs

This Page Blank (uspio)

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 23:45:48 ; Search time 516.754 Seconds
(without alignments)
4708.259 Million cell updates/sec

Title: US-09-551-494-12
Perfect score: 411
Sequence: 1 cctctgcgaattgaactcac.....tgtgcctagtagacagtga 411

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	411	100.0	411	3 AAC62379	Aac62379 Origin of
2	373.8	90.9	6355	3 AAC62372	Aac62372 cDNA sequ
3	373.8	90.9	6355	9 ACC85005	Acc85005 TMV-U2 ge
4	336.2	81.8	5484	8 ADA15011	Ada15011 Tobacco m
5	336.2	81.8	7684	8 ADA15014	Ada15014 Tobacco m
6	336.2	81.8	7684	10 ADB83322	Adb83322 DNA trans
7	336.2	81.8	7685	4 AAD02010	Aad02010 TMV viral
8	336.2	81.8	7685	5 AAD20298	Aad20298 Tobacco m
9	336.2	81.8	7685	5 AAD20295	Aad20295 Tobacco m
10	336.2	81.8	7685	5 AAD20288	Aad20288 Tobacco m
11	336.2	81.8	7685	5 AAD20291	Aad20291 Tobacco m
12	336.2	81.8	7685	5 AAD20296	Aad20296 Tobacco m
13	336.2	81.8	7685	5 AAD20297	Aad20297 Tobacco m
14	336.2	81.8	7685	5 AAD20294	Aad20294 Tobacco m
15	336.2	81.8	7685	6 AAD24472	Aad24472 Tobacco m
16	336.2	81.8	7685	6 AAD24480	Aad24480 Tobacco m
17	336.2	81.8	7685	6 AAD24479	Aad24479 Tobacco m
18	336.2	81.8	7685	6 AAD24478	Aad24478 Tobacco m
19	336.2	81.8	7685	6 AAD24481	Aad24481 Tobacco m
20	336.2	81.8	7685	6 AAD24475	Aad24475 Tobacco m

21	336.2	81.8	7685	6 AAD24482	Aad24482 Tobacco m
22	336.2	81.8	7685	8 ADA15015	Ada15015 Tobacco m
23	336.2	81.8	7685	10 ADB83319	Adb83319 DNA trans
24	336.2	81.8	7685	10 ADB81243	Adb81243 DNA trans
25	336.2	81.8	7685	10 ADB81246	Adb81246 DNA trans
26	336.2	81.8	7686	4 AAD02011	Aad02011 Mutant TM
27	336.2	81.8	7686	5 AAD20292	Aad20292 Tobacco m
28	336.2	81.8	7686	5 AAD20289	Aad20289 Tobacco m
29	336.2	81.8	7686	6 AAD24473	Aad24473 Tobacco m
30	336.2	81.8	7686	6 AAD24476	Aad24476 Tobacco m
31	336.2	81.8	7686	8 ADA15012	Ada15012 Tobacco m
32	336.2	81.8	7686	10 ADB83320	Adb83320 DNA trans
33	336.2	81.8	7686	10 ADB83323	Adb83323 DNA trans
34	336.2	81.8	7686	10 ADB81244	Adb81244 DNA trans
35	336.2	81.8	7686	10 ADB81247	Adb81247 DNA trans
36	336.2	81.8	7687	5 AAD20290	Aad20290 Tobacco m
37	336.2	81.8	7687	6 AAD24474	Aad24474 Tobacco m
38	336.2	81.8	7687	8 ADA15013	Ada15013 Tobacco m
39	336.2	81.8	7687	10 ADB83321	Adb83321 DNA trans
40	336.2	81.8	7687	10 ADB81245	Adb81245 DNA trans
41	336.2	81.8	7688	5 AAD20293	Aad20293 Tobacco m
42	336.2	81.8	7688	6 AAD24477	Aad24477 Tobacco m
43	336.2	81.8	7688	8 ADA15016	Ada15016 Tobacco m
44	336.2	81.8	7688	10 ADB83324	Adb83324 DNA trans
45	336.2	81.8	7688	10 ADB81248	Adb81248 DNA trans

ALIGNMENTS

RESULT 1
AAC62379
ID AAC62379 standard; DNA; 411 BP.

AC AAC62379;

XX 19-MAR-2001 (first entry)

XX Origin of assembly (OAS) of a tobacco mosaic virus (TMV)-U2.

XX Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
KW Gossypium; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
KW tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus;
KW origin of assembly; ss.

XX Tobacco mosaic virus.

XX WO200063397-A2.

XX 26-OCT-2000.

XX 17-APR-2000; 2000WO-EP003521.

XX 20-APR-1999; 99US-00294022.

XX (AVET) AVENTIS CROPS SCIENCE NV.

XX Meulwaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaiff M;

XX WPI; 2000-687182/67.

XX Identifying and isolating genes involved in determining the trait or
PT phenotype of plant species, by infecting plants with gene silencing
PT constructs targeted to the gene, and identifying plants with altered
PT traits.

XX Example 1; Page 63; 64pp; English.

XX The specification describes a method for isolating genes that determine a
CC trait or phenotype of a plant species. The method comprises identifying a
CC set of nucleic acids of genes correlated with the trait, creating a
CC library of gene silencing constructs in a viral RNA vector, targeting the
CC gene silencing constructs to the nucleic acid set, infecting a collection

XX ACC85005;
XX 13-OCT-2003 (first entry)
XX TMV-U2 genome nucleotide sequence.
XX Inhibitory RNA; viral RNA vector; coat protein; TMV; U2; gene; ds.
XX Tobacco mosaic virus.
XX WO2003052108-A2.
XX 26-JUN-2003.
XX 05-DEC-2002; 2002WO-EP013964.
XX 18-DEC-2001; 2001US-0340488P.
XX (FARB) BAYER BIOSCIENCE NV.
XX Metzlaaff MH, Gossele VML, Meulewaeter F, Fache ICA;
XX WPI; 2003-523529/49.
XX
XX Introducing inhibitory RNA into a plant cell comprises providing a viral
XX RNA vector derived from a satellite RNA virus that encodes a coat
XX protein, and infecting a plant with the viral RNA vector and a
XX corresponding helper virus.
XX Example; Page 79-82; 86pp; English.
XX
XX The invention relates to introducing inhibitory RNA into a plant cell.
XX The method involves providing a viral RNA vector derived from a satellite
XX RNA virus having a sequence that encodes a coat protein, and infecting a
XX plant with the viral RNA vector and a corresponding helper virus. The
XX methods and viral RNA vectors are useful in introducing inhibitory RNA
XX into plant cells. These may be used to determine or validate the function
XX of isolated nucleic acid sequences in plants. The present sequence
XX represents the nucleotide sequence of the genome of tobacco mosaic virus
XX (TMV)-U2
XX
XX Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;
Query Match 90.9%; Score 373.8; DB 9; Length 6355;
Best Local Similarity 94.6%; Pred. No. 1.6e-89;
Matches 387; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 3 CTCGCCAATTCGAATCACTGAAAAAGTTGTTGATGAGTTCGTAGATGAAGTACCGATGGC 62
Db 5441 CTCGCCAATTCGAATCACTGAAAAAGTTGTTGAGGAGTTCGTGATGAAGTACCAATGGC 5500
Qy 63 TGTGAATCTGAAGGTTCCGGAAAAACAAAAAGAGAGTGTAGTGAATATGTTAATAA 122
Db 5501 TGTGAATCTGAAGGTTCCGGAAAAACAAAAAGAGATGTAGTGAATATGTTAATAA 5560
Qy 123 TAAGAAATAATAATAGTGTGAAGAGGTTTGAAGTTTCAGGAATTTGAGGATAATGT 182
Db 5561 TAAGAAATAATAATAGTGTGAAGAGGTTTGAAGTTTCAGGAATTTGAGGATAATGT 5620
Qy 183 AAGTGATGACGAGTCTATCGCGTTCATCGAGTACGTTTTTAATCAATATGCTTATACAATC 242
Db 5621 AAGTGATGACGAGTCTATCGCGTTCATCGAGTACGTTTTTAATCAATATGCTTATACAATC 5680
Qy 243 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGTTATGAGATCTCTGTCAGCTGATC 302
Db 5681 AACTCTCGAGCCAAATTTGTTTACTTAATCTCTCCGTTACGCGATCTCTGTCAGCTGATC 5740
Qy 303 AATCTGTGTACAAATAGTAAAGTAAACCAAGTTTCAACGCAACAGCTTAGGACACAGTC 362
Db 5741 AATCTGTGTACAAATAGTAAAGTAAACCAAGTTTCAACGCAACAGCTTAGGACACAGTC 5800
Qy 363 CAACAGCAATTTGCGGATGCGCTGGAAACCTGTGCTAGTATGACAGTGA 411

Db 5801 CAACAGCAATTTGCGGATGCGCTGGAAACCTGTGCTAGTATGACAGTGA 5849
RESULT 4
ADA15011
ID ADA15011 standard; DNA; 5484 BP.
XX AC ADA15011;
XX 06-NOV-2003 (first entry)
XX Tobacco mosaic virus plasmid pBT130BGFPc3 (p1037) .
XX uncapped RNA molecule; plant; transgenic; pest resistant;
XX pathogen resistant; herbicide tolerant; modified growth habit;
XX modified metabolic characteristic; ds.
XX Synthetic.
XX OS Tobacco mosaic virus.
XX US2002164803-A1.
XX 07-NOV-2002.
XX 07-SEP-2001; 2001US-00949317.
XX 16-JAN-1998; 98US-00008186.
XX 15-JAN-1999; 99US-00232170.
XX 21-JUN-1999; 99US-00359301.
XX 11-JUL-1999; 99US-00359305.
XX 11-FEB-2000; 2000US-00502710.
XX (LIND/) LINDBO J A.
XX (POGU/) POGUE G P.
XX (TURP/) TURPEN T H.
XX Lindbo JA, Pogue GP, Turpen TH;
XX WPI; 2003-220044/21.
XX
XX New uncapped RNA molecule of a positive strand replicating RNA virus,
XX useful as RNA transformation vectors for producing phenotypically
XX transformed plants that are e.g. pest or pathogen resistant, or herbicide
XX tolerant.
XX Disclosure; Fig 1; 37pp; English.
XX
XX The invention relates to an uncapped RNA molecule of a single-component
XX single-stranded (+) sense RNA virus, which is capable of infecting a host
XX plant cell. The uncapped RNA molecule comprises a viral replication
XX element, an exogenous RNA segment, and no base, a single base or a
XX sequence of bases located at the 5' terminus of the viral sequence. Also
XX claimed is a method for modifying a host plant cell phenotypically by
XX introducing into the cell the uncapped RNA molecule, where the exogenous
XX RNA segment confers a detectable trait in the host cell, thus modifying
XX the host cell and a DNA transcription vector comprising cDNA having one
XX strand complementary to the uncapped RNA molecule capable of infecting a
XX host plant cell. The uncapped RNA molecule is useful as RNA
XX transformation vectors for modifying a plant host cell. In particular,
XX the uncapped RNA molecule is useful for producing phenotypically
XX transformed plants under field or greenhouse growth conditions to produce
XX plants that are e.g. pest resistant, pathogen resistant, herbicide
XX tolerant, or with modified growth habit and modified metabolic
XX characteristics (e.g. production of commercially useful peptides or
XX pharmaceuticals in plants). The present sequence represents the tobacco
XX mosaic virus plasmid pBT130BGFPc3 (p1037) .
XX
XX Sequence 5484 BP; 1639 A; 1012 C; 1303 G; 1530 T; 0 U; 0 Other;
Query Match 81.8%; Score 336.2; DB 8; Length 5484;
Best Local Similarity 98.9%; Pred. No. 1.7e-79;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
Qy 60 GGCTGTGAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGGTAGTAAATGTTA 118
Db 4480 GGCTGTGAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGGTAGTAAATGTTA 4539
Qy 119 ATAATAAGAAAATAAATAAGTGTGAAGAGGGTTTGAAAAGTTGAGGAAATTCAGGATA 178
Db 4540 ATAATAAGAAAATAAATAAGTGTGAAGAGGGTTTGAAAAGTTGAGGAAATTCAGGATA 4599
Qy 179 ATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCGCTTATAC 238
Db 4600 ATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCGCTTATAC 4659
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGTACCTGTCGAGCT 298
Db 4660 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGTACCTGTCGAGCT 4719
Qy 299 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAAAACGCAACAAAGCTAGGACAAC 358
Db 4720 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAAAACGCAACAAAGCTAGGACAAC 4779
Qy 359 AGTCAACAGCAATTTGCGGATGCTCGAAAACCTGTGCTTAGTATGACAGTGA 411
Db 4780 AGTCAACAGCAATTTGCGGATGCTCGAAAACCTGTGCTTAGTATGACAGTGA 4832

RESULT 5
ADA15014
ID ADA15014 standard; DNA; 7684 BP.
AC ADA15014;
XX
DT 06-NOV-2003 (first entry)
DE Tobacco mosaic virus plasmid pBT11056.
KW uncapped RNA molecule; plant; transgenic; pest resistant;
KW pathogen resistant; herbicide tolerant; modified growth habit;
KW modified metabolic characteristic; db.
XX
OS Synthetic.
OS Tobacco mosaic virus.
XX
PN US2002164803-A1.
XX
PD 07-NOV-2002.
XX
PF 07-SEP-2001; 2001US-00949317.
XX
PR 16-JAN-1998; 98US-00008186.
PR 15-JAN-1999; 99US-00232170.
PR 21-JUL-1999; 99US-00359301.
PR 21-JUL-1999; 99US-00359305.
PR 11-FEB-2000; 2000US-00502710.
XX
PA (LIND/) LINDBO J A.
PA (POGU/) POGUE G P.
PA (TURP/) TURPEN T H.
XX
PI Lindbo JA, Pogue GP, Turpen TH;
XX
DR WPI; 2003-220044/21.
XX
PT New uncapped RNA molecule of a positive strand replicating RNA virus,
PT useful as RNA transformation vectors for producing phenotypically
PT transformed plants that are e.g. pest or pathogen resistant, or herbicide
PT tolerant.
XX
PS Disclosure; Fig 4; 37pp; English.
XX
CC The invention relates to an uncapped RNA molecule of a single-component
CC single-stranded (+) sense RNA virus, which is capable of infecting a host
CC plant cell. The uncapped RNA molecule comprises a viral replication
```

```
CC element, an exogenous RNA segment, and no base, a single base or a
CC sequence of bases located at the 5' terminus of the viral sequence. Also
CC claimed is a method for modifying a host plant cell phenotypically by
CC introducing into the cell the uncapped RNA molecule, where the exogenous
CC RNA segment confers a detectable trait in the host cell, thus modifying
CC the host cell and a DNA transcription vector comprising cDNA having one
CC strand complementary to the uncapped RNA molecule capable of infecting a
CC host plant cell. The uncapped RNA molecule is useful as RNA
CC transformation vectors for modifying a plant host cell. In particular,
CC the uncapped RNA molecule is useful for producing phenotypically
CC transformed plants under field or greenhouse growth conditions to produce
CC plants that are e.g. pest resistant, pathogen resistant, herbicide
CC tolerant, or with modified growth habit and modified metabolic
CC characteristics (e.g. production of commercially useful peptides or
CC pharmaceuticals in plants). The present sequence represents the tobacco
CC mosaic virus plasmid pBT11056.
XX
SQ Sequence 7684 BP; 2277 A; 1463 C; 1823 G; 2121 T; 0 U; 0 Other;

Query Match 81.8%; Score 336.2; DB 8; Length 7684;
Best Local Similarity 98.9%; Pred. No. 1.9e-79;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GGCTGTGAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGGTAGTAAATGTTA 118
Db 6680 GGCTGTGAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGGTAGTAAATGTTA 6739
Qy 119 ATAATAAGAAAATAAATAAGTGTGAAGAGGGTTTGAAAAGTTGAGGAAATTCAGGATA 178
Db 6740 ATAATAAGAAAATAAATAAGTGTGAAGAGGGTTTGAAAAGTTGAGGAAATTCAGGATA 6799
Qy 179 ATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCGCTTATAC 238
Db 6800 ATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCGCTTATAC 6859
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGTACCTGTCGAGCT 298
Db 6860 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGTACCTGTCGAGCT 6919
Qy 299 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAAAACGCAACAAAGCTAGGACAAC 358
Db 6920 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAAAACGCAACAAAGCTAGGACAAC 6979
Qy 359 AGTCAACAGCAATTTGCGGATGCTCGAAAACCTGTGCTTAGTATGACAGTGA 411
Db 6980 AGTCAACAGCAATTTGCGGATGCTCGAAAACCTGTGCTTAGTATGACAGTGA 7032

RESULT 6
ADB83322
ID ADB83322 standard; DNA; 7684 BP.
XX
AC ADB83322;
XX
DT 04-DEC-2003 (first entry)
XX
DE DNA transcription vector pBT11056.
XX
KW ds; DNA transcription vector; capped RNA; plant cell phenotype.
OS Unidentified.
OS Enterobacteria phage T7.
OS Tobacco mosaic virus.
XX
PN US2002168769-A1.
XX
PD 14-NOV-2002.
XX
PF 07-SEP-2001; 2001US-00949316.
XX
PR 16-JAN-1998; 98US-00008186.
PR 15-JAN-1999; 99US-00232170.
PR 21-JUL-1999; 99US-00359301.
```

```
PR 21-JUL-1999; 99US-00359305.
PR 11-FEB-2000; 2000US-00502711.
PA (LIND/) LINDBO J A.
PA (POGU/) POGUE G P.
PA (TURP/) TURPEN T H.
XX
PI Lindbo JA, Pogue GP, Turpen TH;
XX
XX WPI; 2003-719986/68.
XX
XX DNA transcription vector for modifying a plant cell, comprises a
PT complementary strand to a capped RNA molecule having a cis-acting viral
PT replication element, an intervening base sequence and an exogenous RNA
PT segment.
XX
XX Disclosure; Fig 4; 37pp; English.
XX
XX The invention relates to a DNA transcription vector with a complementary
CC strand to a capped RNA infecting a host plant cell, the molecule having a
CC viral sequence with a cis-acting viral replication element, plus a sense
CC RNA plant virus, intervening base sequences between a cap and 5' terminus
CC viral sequence and an exogenous RNA segment and the molecule replicating
CC without trans-acting viral replication element. The capped RNA molecule
CC is useful for modifying a host plant cell phenotypically, which involves
CC introducing the capped RNA molecule into the cell. The exogenous RNA
CC segment of the capped RNA molecule confers a detectable trait in the host
CC plant cell, thus modifying the host plant cell. The present sequence
CC represents the DNA transcription vector pBT11056.
XX
XX Sequence 7684 BP; 2277 A; 1463 C; 1823 G; 2121 T; 0 U; 0 Other;
SQ
Query Match 81.8%; Score 336.2; DB 10; Length 7684;
Best Local Similarity 98.9%; Pred. No. 1.9e-79;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 60 GCCTGTGAACTCG-AAAGGTTCCGAAACAAAAAGAGAGTGTAGTAAATAGTTA 118
Db 6680 GCCTGTGAACTCGAAAGGTTCCGAAACAAAAAGAGAGTGTAGTAAATAGTTA 6739
QY 119 ATATAGAAATAAATAATAGTGTAGTAAAGGTTTGAAGTTGAGGAATAGGATA 178
Db 6740 ATATAGAAATAAATAATAGTGTAGTAAAGGTTTGAAGTTGAGGAATAGGATA 6799
QY 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTATAC 238
Db 6800 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTATAC 6859
QY 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAAGTTCCGCTTATGAGATCCTGTGCAGCT 298
Db 6860 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAAGTTCCGCTTATGAGATCCTGTGCAGCT 6919
QY 299 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAACGCGAACAAGCTAGGACAAAC 358
Db 6920 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAACGCGAACAAGCTAGGACAAAC 6979
QY 359 AGTCCAACAGCAATTTGCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 411
Db 6980 AGTCCAACAGCAATTTGCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 7032
RESULT 7
ID AAD02010 standard; DNA; 7685 BP.
XX
AC AAD02010;
XX
DT 26-MAR-2001 (first entry)
XX
DE TMV viral vector, BSG1037 DNA encoding wild type viral movement protein.
XX TMV; tobacco; viral movement protein; 126/183kDa replicase protein;
KW tobacco mosaic viral vector; genetic manipulation; transgenic plant;
KW
```

```
KW BSG1037; ds.
XX Tobacco mosaic virus.
XX
XX Key Location/Qualifiers
XX CDS 4903..5709
XX /*tag= a
XX /product= "TMV wild type viral movement protein"
XX /note= "The coding sequence is specifically claimed as
XX SEQ ID NO: 3 in claim 3"
XX
XX WO200066743-A2.
XX
XX 09-NOV-2000.
XX
XX 04-MAY-2000; 2000WO-US012380.
XX
XX 04-MAY-1999; 99US-0132697P.
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX Fitzmaurice WP, Pogue GP, Lindbo JA;
XX WPI; 2001-031757/04.
XX P-PSDB; AAY71917.
XX Novel nucleic acid encoding altered viral movement protein and altered
XX 126/183 kDa replicase complex for genetic manipulation of plants.
XX Disclosure; Fig 5; 48pp; English.
XX
XX The present DNA sequence is the Tobacco mosaic viral (TMV) vector
CC BSG1037, encoding wild type viral movement protein. The invention relates
CC to a viral vector encoding an altered viral movement protein and
CC 126/183kDa replicase proteins, with enhanced ability to stabilise a
CC transgene contained in a viral vector. These vectors are useful for the
CC genetic manipulation of plants. Recombinant vectors and altered proteins
CC are useful for enhancing the stabilisation of a transgene which is
CC preferably a non-viral transgene encoding protein like membrane protein,
CC cytosolic protein, secreted protein, nuclear protein or chaperon protein.
CC The invention further provides a transgenic plant comprising the viral
CC vector, for e.g., tobacco plants like Nicotiana benthamiana and Nicotiana
CC tabacum. These transgenic plants are useful for the production of desired
CC proteins, and as test systems for analysis of the biological functions of
CC a gene
XX
XX Sequence 7685 BP; 2276 A; 1464 C; 1825 G; 2120 T; 0 U; 0 Other;
SQ
Query Match 81.8%; Score 336.2; DB 4; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.9e-79;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 60 GCCTGTGAACTCG-AAAGGTTCCGAAACAAAAAGAGAGTGTAGTAAATAGTTA 118
Db 6681 GCCTGTGAACTCGAAAGGTTCCGAAACAAAAAGAGAGTGTAGTAAATAGTTA 6740
QY 119 ATATAGAAATAAATAATAGTGTAGTAAAGGTTTGAAGTTGAGGAATAGGATA 178
Db 6741 ATATAGAAATAAATAATAGTGTAGTAAAGGTTTGAAGTTGAGGAATAGGATA 6800
QY 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTATAC 6860
QY 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAAGTTCCGCTTATGAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAAGTTCCGCTTATGAGATCCTGTGCAGCT 6920
QY 299 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAACGCGAACAAGCTAGGACAAAC 358
Db 6921 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAACGCGAACAAGCTAGGACAAAC 6980
QY 359 AGTCCAACAGCAATTTGCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 411
```

```
Db      6981 AGTCCAAACAGCAATTTGGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 7033
|||||
RESULT 8
AAD20298
ID      AAD20298 standard; DNA; 7685 BP.
XX
XX
AC      AAD20298;
XX
XX      03-JAN-2002 (first entry)
XX
DE      Tobacco mosaic virus (TMV) pBT1 30BGFPc3 vector DNA mutant, G5402A.
XX
XX      Tobacco mosaic virus; TMV; RNA viral vector; transgenic plant; mutant;
KW      ds.
XX
XX      Tobacco mosaic virus.
OS
OS      Synthetic.
XX
XX      Key      Location/Qualifiers
FH      mutation replace(5402, G)
FT      /*tag= a
XX
XX      US6300133-B1.
XX
XX      09-OCT-2001.
XX
XX      11-FEB-2000; 2000US-00502710.
XX
XX      16-JAN-1998; 98US-00008186.
XX      15-JAN-1999; 99US-00232170.
XX      21-JUL-1999; 99US-00359301.
XX      21-JUL-1999; 99US-00359305.
XX
XX      (LARG-) LARGE SCALE BIOLOGY CORP.
XX
XX      Lindbo JA, Pogue GP, Turpen TH;
XX
XX      WPI; 2001-624664/72.
XX
XX      Uncapped RNA molecule of a single-stranded (+) sense RNA virus for of
XX      infecting a host plant cell comprises an exogenous RNA segment capable of
XX      expressing its function.
XX
XX      Example 1; Col; 61pp; English.
XX
XX      The invention relates to the filed of plant viruses, more particularly to
XX      plus strand RNA viral vector for transformation of a host organism with a
XX      foreign RNA and expression of it. The foreign RNA is inserted into an
XX      infective RNA viral segment containing cis-acting viral replication
XX      elements and allowed to infect the host organism. The RNA vector is
XX      modified to obtain infectivity by not incorporating a cap at the 5' end
XX      of the genome. The modified RNA is able to tolerate the exogenous RNA
XX      segment without disrupting the replication of the modified RNA, in the
XX      absence of a trans-acting viral replication element in a single component
XX      plant virus host cell. The uncapped RNA molecule of a single-stranded (+)
XX      sense RNA virus is capable of infecting a host plant cell. This is useful
XX      for the production of transgenic plants and the production of transgenic
XX      cells. The present sequence is Tobacco mosaic virus (TMV) pBT1 30BGFPc3
XX      vector mutant DNA. Note: This sequence is not shown in the specification,
XX      but it is derived from TMV pBT1 30BGFPc3 vector DNA shown in fig 1
XX      (AAD20288)
XX
XX      Sequence 7685 BP; 2277 A; 1464 C; 1824 G; 2120 T; 0 U; 0 Other;
XX
XX      Query Match      81.8%; Score 336.2; DB 5; Length 7685;
XX      Best Local Similarity 98.9%; Pred.No.1.9e-79;
XX      Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
XX
XX      60 GCCTGTGAAATCG-AAAGGTTTCGGAAACAAAAGAGAGTGTGTAGTAAATAGTTA 118
XX      6681 GCCTGTGAAATCGAAAGGTTTCGGAAACAAAAGAGAGTGTGTAGTAAATAGTTA 6740
```

```
QY      119 ATATAAGAAAAATAATAATAGTGGTAAAGGGTTTGAAAGTTGAGAAATTGAGGATA 178
|||||
Db      6741 ATATAAGAAAAATAATAATAGTGGTAAAGGGTTTGAAAGTTGAGAAATTGAGGATA 6800
|||||
QY      179 ATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCCCTTATAC 238
|||||
Db      6801 ATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCCCTTATAC 6860
|||||
QY      239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGCGAGATCCTGTGCGAGCT 298
|||||
Db      6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGCGAGATCCTGTGCGAGCT 6920
|||||
QY      299 GATCAATCTGTGTACAAATGCAATAGTAACCAAGTTTCAACGCAACAAGCTAGACAAAC 358
|||||
Db      6921 GATCAATCTGTGTACAAATGCAATAGTAACCAAGTTTCAACGCAACAAGCTAGACAAAC 6980
|||||
QY      359 AGTCCAAACAGCAATTTGCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 411
|||||
Db      6981 AGTCCAAACAGCAATTTGCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 7033
|||||
RESULT 9
AAD20295
ID      AAD20295 standard; DNA; 7685 BP.
XX
XX      AC      AAD20295;
XX
XX      03-JAN-2002 (first entry)
XX
XX      Tobacco mosaic virus (TMV) pBT1 30BGFPc3 vector DNA mutant, A2382G.
XX
XX      Tobacco mosaic virus; TMV; RNA viral vector; transgenic plant; mutant;
KW      ds.
XX
XX      Tobacco mosaic virus.
OS
OS      Synthetic.
XX
XX      Key      Location/Qualifiers
FH      mutation replace(2382, A)
FT      /*tag= a
XX
XX      US6300133-B1.
XX
XX      09-OCT-2001.
XX
XX      11-FEB-2000; 2000US-00502710.
XX
XX      16-JAN-1998; 98US-00008186.
XX      15-JAN-1999; 99US-00232170.
XX      21-JUL-1999; 99US-00359301.
XX      21-JUL-1999; 99US-00359305.
XX
XX      (LARG-) LARGE SCALE BIOLOGY CORP.
XX
XX      Lindbo JA, Pogue GP, Turpen TH;
XX
XX      WPI; 2001-624664/72.
XX
XX      Uncapped RNA molecule of a single-stranded (+) sense RNA virus for of
XX      infecting a host plant cell comprises an exogenous RNA segment capable of
XX      expressing its function.
XX
XX      Example 1; Col; 61pp; English.
XX
XX      The invention relates to the filed of plant viruses, more particularly to
XX      plus strand RNA viral vector for transformation of a host organism with a
XX      foreign RNA and expression of it. The foreign RNA is inserted into an
XX      infective RNA viral segment containing cis-acting viral replication
XX      elements and allowed to infect the host organism. The RNA vector is
XX      modified to obtain infectivity by not incorporating a cap at the 5' end
XX      of the genome. The modified RNA is able to tolerate the exogenous RNA
XX      segment without disrupting the replication of the modified RNA, in the
XX      absence of a trans-acting viral replication element in a single component
XX      plant virus host cell. The uncapped RNA molecule of a single-stranded (+)
XX      sense RNA virus is capable of infecting a host plant cell. This is useful
XX      for the production of transgenic plants and the production of transgenic
XX      cells. The present sequence is Tobacco mosaic virus (TMV) pBT1 30BGFPc3
XX      vector mutant DNA. Note: This sequence is not shown in the specification,
XX      but it is derived from TMV pBT1 30BGFPc3 vector DNA shown in fig 1
XX      (AAD20288)
XX
XX      Sequence 7685 BP; 2277 A; 1464 C; 1824 G; 2120 T; 0 U; 0 Other;
XX
XX      Query Match      81.8%; Score 336.2; DB 5; Length 7685;
XX      Best Local Similarity 98.9%; Pred.No.1.9e-79;
XX      Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
XX
XX      60 GCCTGTGAAATCG-AAAGGTTTCGGAAACAAAAGAGAGTGTGTAGTAAATAGTTA 118
XX      6681 GCCTGTGAAATCGAAAGGTTTCGGAAACAAAAGAGAGTGTGTAGTAAATAGTTA 6740
```

CC	absence of a trans-acting viral replication element in a single component
CC	plant virus host cell. The uncapped RNA molecule of a single-stranded (+)
CC	sense RNA virus is capable of infecting a host plant cell. This is useful
CC	for the production of transgenic plants and the production of transgenic
CC	cells. The present sequence is Tobacco mosaic virus (TMV) pBT1 30BGFPc3
CC	vector mutant DNA. Note: This sequence is not shown in the specification,
CC	but it is derived from TMV pBT1 30BGFPc3 vector DNA shown in fig 1
CC	(AAD20288)
XX	
SQ	Sequence 7685 BP; 2275 A; 1464 C; 1826 G; 2120 T; 0 U; 0 Other;
	Query Match 81.8%; Score 336.2; DB 5; Length 7685;
	Best Local Similarity 98.9%; Pred. No. 1.9e-79;
	Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy	60 GCCTGTGAAACTCG-AAAGGTTCCGGAAACAAAAAAAGAGAGTGGTAAATGTTA 118
Db	6681 GCCTGTGAAACTCGAAAGGTTCCGGAAACAAAAAAAGAGAGTGGTAAATGTTA 6740
Qy	119 ATATATAGAAAAATAAATAATAGTCGTAAAGAGGCTTTGAAAGTTGAGGAATTCAGGATA 178
Db	6741 ATATATAGAAAAATAAATAATAGTCGTAAAGAGGCTTTGAAAGTTGAGGAATTCAGGATA 6800
Qy	179 ATGTAAGTGATGACGAGCTCTATCCGCTCATCGAGTACGTTTTTAATCAATATCGCTTATAC 238
Db	6801 ATGTAAGTGATGACGAGCTCTATCCGCTCATCGAGTACGTTTTTAATCAATATCGCTTATAC 6860
Qy	239 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATCCAGATCCTGTGCGAGCT 298
Db	6861 AATCAACTCTCCGAGCCAAATTGTTTACTTAAAGTTCCGCTTATCCAGATCCTGTGCGAGCT 6920
Qy	299 GATCAATCTGTGTACAAATGATTAAGTAAACAGTTTCAAACGCCAAACAAGCTAGGACAAC 358
Db	6921 GATCAATCTGTGTACAAATGATTAAGTAAACAGTTTCAAACGCCAAACAAGCTAGGACAAC 6980
Qy	359 AGTCCAAACAGCAATTTCCGGATGCTCGAAACCTGTGCTAGTAGACAGTGA 411
Db	6981 AGTCCAAACAGCAATTTCCGGATGCTCGAAACCTGTGCTAGTAGACAGTGA 7033

XX	RESULT 10
XX	AAD20288
XX	ID AAD20288 standard; DNA; 7685 BP.
XX	AC AAD20288;
XX	DT 03-JAN-2002 (first entry)
XX	DE Tobacco mosaic virus (TMV) pBT1 30BGFPc3 vector DNA.
XX	XX Tobacco mosaic virus; TMV; RNA viral vector; transgenic plant; ds.
XX	OS Tobacco mosaic virus.
XX	PN US6300133-B1.
XX	PD 09-OCT-2001.
XX	PF 11-FEB-2000; 2000US-00502710.
XX	PR 16-JAN-1998; 98US-00008186.
XX	PR 15-JAN-1999; 99US-00232170.
XX	PR 21-JUL-1999; 99US-00359301.
XX	PR 21-JUL-1999; 99US-00359305.
XX	PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX	PI Lindbo JA, Pogue GP, Turpen TH;
XX	PR WPI; 2001-624664/72.
XX	DR
XX	PT
XX	PT Uncapped RNA molecule of a single-stranded (+) sense RNA virus for of
XX	PT infecting a host plant cell comprises an exogenous RNA segment capable of

PT	expressing its function.
XX	
PS	Disclosure; Fig 1; 6lpp; English.
XX	
CC	The invention relates to the filed of plant viruses, more particularly to a plus strand RNA viral vector for transformation of a host organism with a foreign RNA and expression of it. The foreign RNA is inserted into an infective RNA viral segment containing cis-acting viral replication elements and allowed to infect the host organism. The RNA vector is modified to obtain infectivity by not incorporating a cap at the 5' end of the genome. The modified RNA is able to tolerate the exogenous RNA segment without disrupting the replication of the modified RNA, in the absence of a trans-acting viral replication element in a single component plant virus host cell. The uncapped RNA molecule of a single-stranded (+) sense RNA virus is capable of infecting a host plant cell. This is useful for the production of transgenic plants and the production of transgenic cells. The present sequence is Tobacco mosaic virus (TMV) pBTI 30BGFPc3 vector DNA
CC	
XX	
SQ	Sequence 7685 BP; 2276 A; 1464 C; 1825 G; 2120 T; 0 U; 0 Other;
	Query Match 81.8%; Score 336.2; DB 5; Length 7685;
	Best Local Similarity 98.9%; Pred. No. 1.9e-79;
	Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY	60 GGCTGTGAAACTCG-AAAGGTTCCGGAAAAACAAAAGAGAGTGGTAGGTAATAATGTTA 118
DB	6681 GGCTGTGAAACTCGAAAGGTTCCGGAAAAACAAAAGAGAGTGGTAGGTAATAATGTTA 6740
QY	119 ATAAATGAAGAATAATAATAGTGTTAAGNAGGTTTTGAAGTTGAGGAATTGCAGGATA 178
DB	6741 ATAAATGAAGAATAATAATAGTGTTAAGNAGGTTTTGAAGTTGAGGAATTGCAGGATA 6800
QY	179 ATGTAAGTCATGACGAGTCTATCCGCTCATCGAGTAGCTTTTAAATCAATATGCTTTATC 238
DB	6801 ATGTAAGTCATGACGAGTCTATCCGCTCATCGAGTAGCTTTTAAATCAATATGCTTTATC 6860
QY	239 AATCAAACCTCCGAGCCAATTTGTTTACTTAAGTTCGCTTATGCGAGATCCTGTGCAGCT 298
DB	6861 AATCAAACCTCCGAGCCAATTTGTTTACTTAAGTTCGCTTATGCGAGATCCTGTGCAGCT 6920
QY	299 GATCAATCTGTGTACAAATGCATTAGGTACCAGTTTCAAACGCCAACAGCTAGGACAAC 358
DB	6921 GATCAATCTGTGTACAAATGCATTGGGTAAACCAAGTTTCAAACGCCAACAGCTAGGACAAC 6980
QY	359 AGTCCAAACAGCAATTTGGCGATGCCCTGGAAACCTGTGCCTAGTATGCACGTGA 411
DB	6981 AGTCCAAACAGCAATTTGGCGATGCCCTGGAAACCTGTGCCTAGTATGCACGTGA 7033
RESULT 11	
AAD20291	
ID	AAD20291 standard; DNA; 7685 BP.
XX	
AC	AAD20291;
XX	
DT	03-JAN-2002 (first entry)
XX	
DE	Tobacco mosaic virus (TMV) pBTI1056 vector DNA.
KW	Tobacco mosaic virus; TMV; RNA viral vector; transgenic plant; ds.
OS	Tobacco mosaic virus.
XX	
FN	US6300133-B1.
XX	
PD	09-OCT-2001.
XX	
PF	11-FEB-2000; 2000US-00502710.
XX	
PR	16-JAN-1998; 98US-00008186.
PR	15-JAN-1999; 99US-00232170.
PR	21-JUL-1999; 99US-000359301.

```
PR 21-JUL-1999; 99US-00359305.
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX Lindbo JA, Pogue GP, Turpen TH;
XX WPI; 2001-624664/72.
XX
PT Uncapped RNA molecule of a single-stranded (+) sense RNA virus for of
PT infecting a host plant cell comprises an exogenous RNA segment capable of
PT expressing its function.
XX
XX Disclosure; Fig 4; 61pp; English.
XX
XX The invention relates to the filed of plant viruses, more particularly to
XX plus strand RNA viral vector for transformation of a host organism with a
XX foreign RNA and expression of it. The foreign RNA is inserted into an
XX infective RNA viral segment containing cis-acting viral replication
XX elements and allowed to infect the host organism. The RNA vector is
XX modified to obtain infectivity by not incorporating a cap at the 5' end
XX of the genome. The modified RNA is able to tolerate the exogenous RNA
XX segment without disrupting the replication of the modified RNA, in the
XX absence of a trans-acting viral replication element in a single component
XX plant virus host cell. The uncapped RNA molecule of a single-stranded (+)
XX sense RNA virus is capable of infecting a host plant cell. This is useful
XX for the production of transgenic plants and the production of transgenic
XX cells. The present sequence is Tobacco mosaic virus (TMV) pBT11056 vector
XX DNA
XX
XX Sequence 7685 BP; 2277 A; 1463 C; 1824 G; 2121 T; 0 U; 0 Other;
XX
XX Query Match 81.8%; Score 336.2; DB 5; Length 7685;
XX Best Local Similarity 98.9%; Pred. No. 1.9e-79;
XX Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
XX
Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAACAAAAAGAGAGTGTAGTAAATAGTTA 118
Db 6681 GCCTGTGAAACTCGAAAGGTTCCGGAACAAAAAGAGAGTGTAGTAAATAGTTA 6740
Qy 119 ATAATAAGAAAAATAAATAATAGTGTGAAGAGGTTTGAAAGTTGAGGAAATGAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAATAGTGTGAAGAGGTTTGAAAGTTGAGGAAATGAGGATA 6800
Qy 179 ATGTAAGTGTAGCAGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTTATAC 238
Db 6801 ATGTAAGTGTAGCAGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGATGTAACAGTTTCAAAACGCAACAAAGTAGGACAAAC 358
Db 6921 GATCAATCTGTGTACAAATGATGTAACAGTTTCAAAACGCAACAAAGTAGGACAAAC 6980
Qy 359 AGTCCACAGCAATTTGCGGATGCTGGAACCTGTGCTAGTATGACAGTA 411
Db 6981 AGTCCACAGCAATTTGCGGATGCTGGAACCTGTGCTAGTATGACAGTA 7033
XX
RESULT 12
AAD20296
ID AAD20296 standard; DNA; 7685 BP.
XX
XX AAD20296;
XX
XX 03-JAN-2002 (first entry)
XX
XX Tobacco mosaic virus (TMV) pBT1 30BGFPe3 vector DNA mutant, C5213T.
XX
XX Tobacco mosaic virus; TMV; RNA viral vector; transgenic plant; mutant;
XX ds.
```

```
OS Tobacco mosaic virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX mutation replace(5213, C)
XX /*tag= a
XX
XX US6300133-B1.
XX
XX 09-OCT-2001.
XX
XX 11-FEB-2000; 2000US-00502710.
XX
XX 16-JAN-1998; 98US-00008186.
XX 15-JAN-1999; 99US-00232170.
XX 21-JUL-1999; 99US-00359301.
XX 21-JUL-1999; 99US-00359305.
XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.
XX
XX Lindbo JA, Pogue GP, Turpen TH;
XX WPI; 2001-624664/72.
XX
XX Uncapped RNA molecule of a single-stranded (+) sense RNA virus for of
XX infecting a host plant cell comprises an exogenous RNA segment capable of
XX expressing its function.
XX
XX Example 1; Col; 61pp; English.
XX
XX The invention relates to the filed of plant viruses, more particularly to
XX plus strand RNA viral vector for transformation of a host organism with a
XX foreign RNA and expression of it. The foreign RNA is inserted into an
XX infective RNA viral segment containing cis-acting viral replication
XX elements and allowed to infect the host organism. The RNA vector is
XX modified to obtain infectivity by not incorporating a cap at the 5' end
XX of the genome. The modified RNA is able to tolerate the exogenous RNA
XX segment without disrupting the replication of the modified RNA, in the
XX absence of a trans-acting viral replication element in a single component
XX plant virus host cell. The uncapped RNA molecule of a single-stranded (+)
XX sense RNA virus is capable of infecting a host plant cell. This is useful
XX for the production of transgenic plants and the production of transgenic
XX cells. The present sequence is Tobacco mosaic virus (TMV) pBT1 30BGFPe3
XX vector mutant DNA. Note: This sequence is not shown in the specification,
XX but it is derived from TMV pBT1 30BGFPe3 vector DNA shown in fig 1
XX (AAD20288)
XX
XX Sequence 7685 BP; 2276 A; 1463 C; 1825 G; 2121 T; 0 U; 0 Other;
XX
XX Query Match 81.8%; Score 336.2; DB 5; Length 7685;
XX Best Local Similarity 98.9%; Pred. No. 1.9e-79;
XX Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
XX
Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAACAAAAAGAGAGTGTAGTAAATAGTTA 118
Db 6681 GCCTGTGAAACTCGAAAGGTTCCGGAACAAAAAGAGAGTGTAGTAAATAGTTA 6740
Qy 119 ATAATAAGAAAAATAAATAATAGTGTGAAGAGGTTTGAAAGTTGAGGAAATGAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAATAGTGTGAAGAGGTTTGAAAGTTGAGGAAATGAGGATA 6800
Qy 179 ATGTAAGTGTAGCAGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTTATAC 238
Db 6801 ATGTAAGTGTAGCAGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGATGTAACAGTTTCAAAACGCAACAAAGTAGGACAAAC 358
Db 6921 GATCAATCTGTGTACAAATGATGTAACAGTTTCAAAACGCAACAAAGTAGGACAAAC 6980
XX
```

Qy 359 AGTCCAAACAGCAATTTGGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 13
AAD20297
ID AAD20297 standard; DNA; 7685 BP.
XX AC AAD20297;
XX XX
XX 03-JAN-2002 (first entry)
XX DE Tobacco mosaic virus (TMV) pBTI 30BGFPc3 vector DNA mutant, A5303G.
XX XX
XX KW Tobacco mosaic virus; TMV; RNA viral vector; transgenic plant; mutant;
XX KW ds.
XX OS Tobacco mosaic virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT mutation replace(5303, A)
XX FT /*tag= a
XX PN US6300133-B1.
XX PD 09-OCT-2001.
XX XX
XX PF 11-FEB-2000; 2000US-00502710.
XX PR 16-JAN-1998; 98US-00008186.
XX PR 15-JAN-1999; 99US-00232170.
XX PR 21-JUL-1999; 99US-00359301.
XX PR 21-JUL-1999; 99US-00359305.
XX XX
XX PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX XX
XX PI Lindbo JA, Pogue GP, Turpen TH;
XX XX
XX DR WPI; 2001-624664/72.
XX XX
XX PT Uncapped RNA molecule of a single-stranded (+) sense RNA virus for of
XX PT infecting a host plant cell comprises an exogenous RNA segment capable of
XX PT expressing its function.
XX XX
XX PS Example 1; Col; 6lpp; English.
XX XX
XX CC The invention relates to the filed of plant viruses, more particularly to
XX CC plus strand RNA viral vector for transformation of a host organism with a
XX CC foreign RNA and expression of it. The foreign RNA is inserted into an
XX CC infective RNA viral segment containing cis-acting viral replication
XX CC elements and allowed to infect the host organism. The RNA vector is
XX CC modified to obtain infectivity by not incorporating a cap at the 5' end
XX CC of the genome. The modified RNA is able to tolerate the exogenous RNA
XX CC segment without disrupting the replication of the modified RNA, in the
XX CC absence of a trans-acting viral replication element in a single component
XX CC plant virus host cell. The uncapped RNA molecule of a single-stranded (+)
XX CC sense RNA virus is capable of infecting a host plant cell. This is useful
XX CC for the production of transgenic plants and the production of transgenic
XX CC cells. The present sequence is Tobacco mosaic virus (TMV) pBTI 30BGFPc3
XX CC vector mutant DNA. Note: This sequence is not shown in the specification,
XX CC but it is derived from TMV pBTI 30BGFPc3 vector DNA shown in fig 1
XX CC (AAD20288)
XX XX
XX SQ Sequence 7685 BP; 2275 A; 1464 C; 1826 G; 2120 T; 0 U; 0 Other;

Query Match 81.8%; Score 336.2; DB 5; Length 7685;
Best Local Similarity 98.9%; Pred No. 1.9e-79;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GGCTGTGAACTCG-AAAGGTTCCGAAACAAAGAGAGTGGTAGGTAATAGTTA 118
|||||

Db 6681 GGCTGTGAACTCGAAAGGTTCCGAAACAAAGAGAGTGGTAGGTAATAGTTA 6740
Qy 119 ATAATAAGAAAAATAATAATAGTGTGAAGAGGTTTGAAGTTGAGAAATTCAGAGTA 178
Db 6741 ATAATAAGAAAAATAATAATAGTGTGAAGAGGTTTGAAGTTGAGAAATTCAGAGTA 6800
Qy 179 ATGTAAGTATGACAGAGTCTATCCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTATGACAGAGTCTATCCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAAACAGTTTCAACGCAACAAAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAAACAGTTTCAACGCAACAAAGCTAGGACAAC 6980
Qy 359 AGTCCAAACAGCAATTTGGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 14
AAD20294
ID AAD20294 standard; DNA; 7685 BP.
XX AC AAD20294;
XX XX
XX DT 03-JAN-2002 (first entry)
XX DE Tobacco mosaic virus (TMV) pBTI 30BGFPc3 vector DNA mutant, A1138G.
XX XX
XX KW Tobacco mosaic virus; TMV; RNA viral vector; transgenic plant; mutant;
XX KW ds.
XX OS Tobacco mosaic virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT mutation replace(1138, A)
XX FT /*tag= a
XX PN US6300133-B1.
XX PD 09-OCT-2001.
XX XX
XX PF 11-FEB-2000; 2000US-00502710.
XX PR 16-JAN-1998; 98US-00008186.
XX PR 15-JAN-1999; 99US-00232170.
XX PR 21-JUL-1999; 99US-00359301.
XX PR 21-JUL-1999; 99US-00359305.
XX XX
XX PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX XX
XX PI Lindbo JA, Pogue GP, Turpen TH;
XX XX
XX DR WPI; 2001-624664/72.
XX XX
XX PT Uncapped RNA molecule of a single-stranded (+) sense RNA virus for of
XX PT infecting a host plant cell comprises an exogenous RNA segment capable of
XX PT expressing its function.
XX XX
XX PS Example 1; Col; 6lpp; English.
XX XX
XX CC The invention relates to the filed of plant viruses, more particularly to
XX CC plus strand RNA viral vector for transformation of a host organism with a
XX CC foreign RNA and expression of it. The foreign RNA is inserted into an
XX CC infective RNA viral segment containing cis-acting viral replication
XX CC elements and allowed to infect the host organism. The RNA vector is
XX CC modified to obtain infectivity by not incorporating a cap at the 5' end
XX CC of the genome. The modified RNA is able to tolerate the exogenous RNA
XX CC segment without disrupting the replication of the modified RNA, in the
XX CC absence of a trans-acting viral replication element in a single component
XX CC plant virus host cell. The uncapped RNA molecule of a single-stranded (+)
XX CC sense RNA virus is capable of infecting a host plant cell. This is useful
XX CC for the production of transgenic plants and the production of transgenic
XX CC cells. The present sequence is Tobacco mosaic virus (TMV) pBTI 30BGFPc3
XX CC vector mutant DNA. Note: This sequence is not shown in the specification,
XX CC but it is derived from TMV pBTI 30BGFPc3 vector DNA shown in fig 1
XX CC (AAD20288)
XX XX
XX SQ Sequence 7685 BP; 2275 A; 1464 C; 1826 G; 2120 T; 0 U; 0 Other;

CC segment without disrupting the replication of the modified RNA, in the
CC absence of a trans-acting viral replication element in a single component
CC plant virus host cell. The uncapped RNA molecule of a single-stranded (+)
CC sense RNA virus is capable of infecting a host plant cell. This is useful
CC for the production of transgenic plants and the production of transgenic
CC cells. The present sequence is Tobacco mosaic virus (TMV) pBTI 30BGFPc3
CC vector mutant DNA. Note: This sequence is not shown in the specification,
CC but it is derived from TMV pBTI 30BGFPc3 vector DNA shown in fig 1
CC (A2D20288)

XX SQ Sequence 7685 BP; 2275 A; 1464 C; 1826 G; 2120 T; 0 U; 0 Other;

Query Match 81.8%; Score 336.2; DB 5; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.9e-79;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGGTAAATGTTA 118
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGGTAAATGTTA 6740
Qy 119 ATAATAAGAAATAAATAAGTGTGAAGAGGTTTGAAGTTGAGGAATTCAGGATA 178
Db 6741 ATAATAAGAAATAAATAAGTGTGAAGAGGTTTGAAGTTGAGGAATTCAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGCTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGCTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTAACAATGATAGTAAGTAAAGTTCGTTTAAATCAATATGCCTTATAC 358
Db 6921 GATCAATCTGTGTAACAATGATAGTAAGTAAAGTTCGTTTAAATCAATATGCCTTATAC 6980
Qy 359 AGTCCACAGCAATTTCCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCACAGCAATTTCCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 15

AAD24472
ID AAD24472 standard; DNA; 7685 BP.
XX
AC AAD24472;
XX
DT 07-MAR-2002 (first entry)
XX
DE Tobacco mosaic virus (TMV) pBTI 30BGFPc3 vector DNA.
XX
KW Tobacco mosaic virus; TMV; RNA viral vector; infectivity;
KW pBTI 30BGFPc3 vector; ds.
XX
OS Tobacco mosaic virus.
XX
FN US6300134-B1.
XX
PD 09-OCT-2001.
XX
PF 11-FEB-2000; 2000US-00502711.
XX
PR 16-JAN-1998; 98US-00008186.
PR 15-JAN-1999; 99US-00232170.
PR 21-JUL-1999; 99US-00359301.
PR 21-JUL-1999; 99US-00359305.
XX
FA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
FI Lindbo JA, Pogue GF, Turpen TH;
XX
DR WPI; 2002-033177/04.
XX

PT A capped RNA molecule derived from a plus strand RNA plant virus having
PT extra nucleotides inserted between the cap and the 5 prime terminus,
PT useful as a vector to introduce exogenous RNA into a host plant cell.

XX Claim 15; Fig 1; 61pp; English.

XX The invention relates to the field of plant viruses, more particularly to
CC plus strand RNA viral vector for transformation of a host organism with a
CC foreign RNA and its expression. The foreign RNA is inserted into an
CC infective RNA viral segment containing cis-acting viral replication
CC elements and allowed to infect the host organism. The RNA vector is
CC modified to obtain infectivity by including an intervening sequence
CC between the cap and the 5' end. The modified RNA is able to tolerate the
CC exogenous RNA segment without disrupting the replication of the modified
CC RNA, in the absence of a trans-acting viral replication element in a
CC single component plant virus host cell. The RNA molecule is used to
CC introduce exogenous RNA into a plant cell. The present sequence is
CC Tobacco mosaic virus (TMV) pBTI 30BGFPc3 vector DNA

XX SQ Sequence 7685 BP; 2276 A; 1464 C; 1825 G; 2120 T; 0 U; 0 Other;

Query Match 81.8%; Score 336.2; DB 6; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.9e-79;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGGTAAATGTTA 118
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGGTAAATGTTA 6740
Qy 119 ATAATAAGAAATAAATAAGTGTGAAGAGGTTTGAAGTTGAGGAATTCAGGATA 178
Db 6741 ATAATAAGAAATAAATAAGTGTGAAGAGGTTTGAAGTTGAGGAATTCAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGCTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGCTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTAACAATGATAGTAAGTAAAGTTCGTTTAAATCAATATGCCTTATAC 358
Db 6921 GATCAATCTGTGTAACAATGATAGTAAGTAAAGTTCGTTTAAATCAATATGCCTTATAC 6980
Qy 359 AGTCCACAGCAATTTCCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCACAGCAATTTCCGGATGCTCGAAACCTGTGCTAGTATGACAGTGA 7033

Search completed: September 26, 2005, 01:44:58
Job time : 517.754 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 00:52:04 ; Search time 3328.26 Seconds
(without alignments)
4700.480 Million cell updates/sec

Title: US-09-551-494-12
Perfect score: 411
Sequence: 1 cccctgcgaattgaactcac.....tgtgcctagtagcacagtga 411

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	73.2	17.8	348	4	BM067518 KS08006E1
2	52.6	12.8	636	4	BJ418369 BJ418369
3	51	12.4	589	4	BJ423793 BJ423793
4	51	12.4	592	4	BJ423613 BJ423613
5	51	12.4	604	4	BJ366800 BJ366800
6	51	12.4	605	4	BJ426769 BJ426769
7	51	12.4	607	4	BJ417997 BJ417997
8	51	12.4	626	4	BJ426777 BJ426777
9	47.2	11.5	277	7	T18065
10	46.4	11.3	1204	9	CNS016E2
11	46.2	11.2	494	6	CD197789
12	45.6	11.1	365	4	BM274859
13	45.4	11.0	645	4	BJ404348
14	45.4	11.0	726	6	CB552291
15	45.2	11.0	729	7	CV527869
16	45.2	11.0	764	2	BB623547
17	45	10.9	798	9	AG570527
18	44.8	10.9	330	4	BI507830
19	44.8	10.9	533	5	BQ451797
20	44.8	10.9	594	5	BU498120
21	44.8	10.9	709	5	BU550168
22	44.4	10.8	634	9	CNS021MB
23	44.2	10.8	428	8	AQ013057
24	44.2	10.8	842	9	CG945985

25	44	10.7	612	4	BJ332345
26	44	10.7	911	9	AG031756
27	44	10.7	1140	9	CNS007M9
28	43.8	10.7	488	4	BI670589
29	43.8	10.7	608	8	AA550167
30	43.8	10.7	730	5	EX100679
31	43.8	10.7	2333	3	BC015933
32	43.6	10.6	432	8	BH126495
33	43.6	10.6	605	4	BM167791
34	43.6	10.6	631	3	AY066111
35	43.6	10.6	731	8	AZ197271
36	43.6	10.6	1638	8	CC234209
37	43.4	10.6	414	9	CL964217
38	43.4	10.6	573	2	AW399615
39	43.4	10.6	626	9	EX120494
40	43.4	10.6	683	6	CD486119
41	43.2	10.5	970	9	CL076055
42	43	10.5	253	8	BH657578
43	43	10.5	350	8	AZ213169
44	43	10.5	584	1	AU033688
45	42.8	10.4	447	4	BI815960

ALIGNMENTS

RESULT 1
BM067518/c 348 bp mRNA linear EST 11-SEP-2002
LOCUS BM067518 KS08006E10 KS08 Capsicum annuum cDNA, mRNA sequence.
DEFINITION BM067518
ACCESSION BM067518.1 GI:22787638
VERSION
KEYWORDS
SOURCE
ORGANISM
Capsicum annuum

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 348.
Location/Qualifiers
1. 348
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stage="10 weeks after germination"
/clone_lib="KS08"
/note="vector: pBluescript SK(-)"

Query Match 17.8%; Score 73.2; DB 4; Length 348;
Best Local Similarity 54.4%; Pred. No. 2.4e-08;
Matches 147; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
Qy 7 CCAATTGAACCTCACTGAAAGTGTTCATGAGTTCGTAGTGAAGTACCGATCGCTGTG 66
Db 279 CCGTGAACCTACAGAGCAGTGTTCATGAGTTCGTAGTTCAGTTCCTCAATGCTGAC 220

Qy 67 AAATCTCGAAGGTTCCGGAACAAACAAAAGAGAGTGGTAGTGAATAATGTTAATAATAAG 126
 Db 219 AGATTACGTAATTTCCGAATCAATCCAGAAAGAGAGTAATAGTATGAGTGAAGAGA 160
 Qy 127 AAAATAATAATAGTGGTGAAGAGGTTTGAAGGTGAGGAAATGAGGATAATGTAAGT 186
 Db 159 AATGATAATAAGGGTGTGAATAAGGAAGGAAGCTGTTGTGATAAGGTTAGAAATGGGCAG 100
 Qy 187 GATGACGAGTCTATCGCTCATCGAGTACGTTTAAATCAATATGCCCTATACATCAACT 246
 Db 99 AACTCGGAGTCATCGAGCGCGAGTCTTCTTCGTTTAACTATGGCTTACACAGTTTCCA 40
 Qy 247 CTCGAGCGCAATTTGTTTACTTAAAGTTCCG 276
 Db 39 GTGCCAATCAATTAGTGTATTAGGTTCTG 10

RESULT 2
 BJ418369
 LOCUS
 DEFINITION BJ418369 Dictyostelium discoideum cDNA library, linear EST 10-MAR-2002
 discoideum cDNA clone ddv32c18 5', mRNA sequence.
 ACCESSION
 VERSION BJ418369
 SOURCE
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 REFERENCE 1 (bases 1 to 636)
 AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
 TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES
 source
 1..636
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="ddv32c18"
 /sex="mat A"
 /dev_stage="Growth phase"
 /clone_lib="Dictyostelium discoideum cDNA library, VF"
 ORIGIN
 Query Match 12.8%; Score 52.6; DB 4; Length 636;
 Best Local Similarity 51.5%; Pred. No. 0.006;
 Matches 121; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
 Qy 24 AAAAGTGTTCGATGCTCGTAGATGAAGTACCGATGCGTGTGAACTCGAAGGTTCCG 83
 Db 326 AGAAGGATGGGTAAATTCGAGAGAAAAGAAAAGAACTGGTAACGACACAAACAACAA 385
 Qy 84 GAAACAAAAGAGAGGTTAGGTGAATAATGTTAATAATGAAGAAATAATAATATAGTGG 143
 Db 386 CAAAGGCAACAATAATAATAAGATAATAAAGATAATAAAGATAATAAAGATAATAAGA 445
 Qy 144 TAAGAAGGTTTGAAGTTGAGGAAATTGAGGATAATGTAAGTGTATGACGAGTCTATCGC 203
 Db 446 TAATAAGATAATAAGATGAAGTTTATGCAACAATGAATTCACCAGATTTAAGAA 505
 Qy 204 GTCATCGAGTACGTTTAAATCAATATGCGCTTATCAATCAACTCTCCGAGCCAAT 258
 Db 506 ATATTTAATACAGTTATATCAATTTGGTGTATATATATATATATATATATATATAT 560

RESULT 3
 BJ423793
 LOCUS
 DEFINITION BJ423793 Dictyostelium discoideum cDNA library, linear EST 11-MAR-2002
 discoideum cDNA clone ddv50n09 5', mRNA sequence.
 ACCESSION
 VERSION BJ423793
 SOURCE
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 REFERENCE 1 (bases 1 to 589)
 AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
 TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES
 source
 1..589
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="ddv50n09"
 /sex="mat A"
 /dev_stage="Growth phase"
 /clone_lib="Dictyostelium discoideum cDNA library, VF"
 ORIGIN
 Query Match 12.4%; Score 51; DB 4; Length 589;
 Best Local Similarity 51.1%; Pred. No. 0.016;
 Matches 120; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
 Qy 24 AAAAGTGTTCGATGCTCGTAGATGAAGTACCGATGCGTGTGAACTCGAAGGTTCCG 83
 Db 310 AGAAGGATGGGTAAATTCGAGAGAAAAGAAAAGAACTGGTAACGACACAAACAACAA 369
 Qy 84 GAAACAAAAGAGAGGTTAGTGAATAATGTTAATAATGAAGAAATAATAATATAGTGG 143
 Db 370 CAAAGGCAACAATAATAATAAGATAATAAAGATAATAAAGATAATAAAGATAATAAGA 429
 Qy 144 TAAGAAGGTTTGAAGTTGAGGAAATTGAGGATAATGTAAGTGTATGACGAGTCTATCGC 203
 Db 430 TAATAAGATAATAAGATGAAGTTTATGCAACAATGAATTCACCAGATTTAAGAA 489
 Qy 204 GTCATCGAGTACGTTTAAATCAATATGCGCTTATCAATCAACTCTCCGAGCCAAT 258
 Db 490 ATATTTAATACAGTTATATCAATTTGGTGTATATATATATATATATATATATATAT 544

RESULT 4
 BJ423613
 LOCUS
 DEFINITION BJ423613 Dictyostelium discoideum cDNA library, linear EST 11-MAR-2002
 discoideum cDNA clone ddv49g21 5', mRNA sequence.
 ACCESSION
 VERSION BJ423613
 SOURCE
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 REFERENCE 1 (bases 1 to 592)
 AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
 TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source
1..592
/organism="Dictyostelium discoideum"
/mol_type="mrna"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv4921"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Query Match 12.4%; Score 51; DB 4; Length 592;
Best Local Similarity 51.1%; Pred. No. 0.016;
Matches 120; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 24 AAAAGTTGTTGATGAGTTCGTAGATGAAGTACCGATGCTGTGAAACTCGAAAGGTTCCG 83
DB 313 AGAAGGATGGGTAAATTCGAAAAAAGAAAGAAAGAACTGTTACGCAACAACA 372
QY 84 GAAACAAAAAGAGAGTGGTAGGTAATAATGTTAATAAGAAATAATAATAGTAGTG 143
DB 373 CAAAGGCAACAATAATAAAGATAATAAAGATAATAAAGATAATAAAGATAATAAAGA 432
QY 144 TAAGAGGTTTGAAGTTGAGGAAATTGAGGATAATGTAAGTGAATGACGAGTCTATCGC 203
DB 433 TAATAAAGATAATAAAGATGAAAGTTTTATGCAACAATGAATTCACCAGAAATTAAGAA 492
QY 204 GTCATCGAGTACGTTTTTAATCAATATGCTTATACATCAACTCTCCGAGCCAAT 258
DB 493 ATATTTAATACAGTTATATCAATGGTGTATATATCACTTTTAGACCAAAAT 547

RESULT 5

LOCUS BJ366800 604 bp mRNA linear EST 08-MAR-2002
DEFINITION BJ366800 Dictyostelium discoideum cDNA library, CF Dictyostelium
discoideum cDNA clone ddc40e10 5', mRNA sequence.

ACCESSION

BJ366800

VERSION

BJ366800.1 GI:19276102

KEYWORDS

EST.

SOURCE

Dictyostelium discoideum

ORGANISM

Dictyostelium discoideum

REFERENCE

1 (bases 1 to 604)

AUTHORS

Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

TITLE

Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

JOURNAL

Full length cDNA of Dictyostelium discoideum at the culmination

COMMENT

stage
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source
1..604
/organism="Dictyostelium discoideum"
/mol_type="mrna"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddc40e10"
/sex="mat A"
/dev_stage="Culmination stage"
/clone_lib="Dictyostelium discoideum cDNA library, CF"

ORIGIN

Query Match 12.4%; Score 51; DB 4; Length 604;
Best Local Similarity 51.1%; Pred. No. 0.016;
Matches 120; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 24 AAAAGTTGTTGATGAGTTCGTAGATGAAGTACCGATGCTGTGAAACTCGAAAGGTTCCG 83
DB 326 AGAAGGATGGGTAAATTCGAAAAAAGAAAGAAAGAACTGTTACGCAACAACA 385
QY 84 GAAACAAAAAGAGAGTGGTAGGTAATAATGTTAATAAAGAAAAATAATAATAGTAGTG 143
DB 386 CAAAGGCAACAATAATAAAGATAATAAAGATAATAAAGATAATAAAGATAATAAAGA 445
QY 144 TAAGAGGTTTGAAGTTGAGGAAATTGAGGATAATGTAAGTGAATGACGAGTCTATCGC 203
DB 446 TAATAAAGATAATAAAGATGAAAGTTTTATGCAACAATGAATTCACCAGAAATTAAGAA 505
QY 204 GTCATCGAGTACGTTTTTAATCAATATGCTTATACATCAACTCTCCGAGCCAAT 258
DB 506 ATATTTAATACAGTTATATCAATGGTGTATATATCACTTTTAGACCAAAAT 560

RESULT 6

BJ426769

LOCUS

BJ426769 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv60n04 5', mRNA sequence.

ACCESSION

BJ426769

VERSION

BJ426769.1 GI:19344205

KEYWORDS

EST.

SOURCE

Dictyostelium discoideum

ORGANISM

Dictyostelium discoideum

REFERENCE

1 (bases 1 to 605)

AUTHORS

Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

TITLE

Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

JOURNAL

Full length cDNA of Dictyostelium discoideum at the vegetative

COMMENT

stage
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source

1..605
/organism="Dictyostelium discoideum"
/mol_type="mrna"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv60n04"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Query Match 12.4%; Score 51; DB 4; Length 605;
Best Local Similarity 51.1%; Pred. No. 0.016;
Matches 120; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 24 AAAAGTTGTTGATGAGTTCGTAGATGAAGTACCGATGCTGTGAAACTCGAAAGGTTCCG 83
DB 326 AGAAGGATGGGTAAATTCGAAAAAAGAAAGAAAGAACTGTTACGCAACAACA 385
QY 84 GAAACAAAAAGAGAGTGGTAGGTAATAATGTTAATAAAGAAAAATAATAATAGTAGTG 143
DB 386 CAAAGGCAACAATAATAAAGATAATAAAGATAATAAAGATAATAAAGATAATAAAGA 445
QY 144 TAAGAGGTTTGAAGTTGAGGAAATTGAGGATAATGTAAGTGAATGACGAGTCTATCGC 203
DB 446 TAATAAAGATAATAAAGATGAAAGTTTTATGCAACAATGAATTCACCAGAAATTAAGAA 505

[illegible]

3

[illegible]

RESULT 11	
CD197789	CD197789
LOCUS	MS1-0095P-V361-H03-U.B MS1-0095 Schistosoma mansoni cDNA Clone
DEFINITION	MS1-0095P-V361-H03.U.B, mRNA sequence.
ACCESSION	CD197789
VERSION	CD197789.1 GI:34727642
KEYWORDS	EST.
SOURCE	Schistosoma mansoni
ORGANISM	Schistosoma mansoni
	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;

REFERENCES

1. Iwasaki, T. *et al.*, 1957.

AUTHORS

Iwajowski-Almeida, S., DeMarco, R., Martins, E. A. L., Guimaraes, P. E. M., Ofati, E. P. B., Paquola A. C. M., Piazza, J. P., Nishiyama, M. Y. Jr., Kitajima, J. P., Adamson, R. E., Ashton, P. D., Bonaldo, M. F., Coulson, P. S., Dillon, G. P., Farias, L. P., Gregorio, S. P., Ho, P. L., Leite, R. A., Malaquias, L. C. C., Marques, R. C. P., Miyasato, P. A., Nascimento, A. L. T. O., Ohlweiler, F. P., Reis, E. M., Ribeiro, M. A.,

TITLE	JOURNAL	Nat. Genet.	35 (2), 148-157 (2003)
transcriptome analysis of the acclimated human parasite <i>Schistosoma mansoni</i>	MEDLINE	22879926	
	PUBMED	12973350	
COMMENT	Contact:	Dr. Sergio Verjovski-Almeida	

Departamento de Bioquímica
Instituto de Química - Universidade de São Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,
Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verto@iq.usp.br

This sequence was derived from the FAPESP Schistosoma mansoni EST
Genome Project. All sequences in the project were assembled and
annotated. This entry and all the assembled sequences can be seen
in the following URL <http://bioinfo.iq.usp.br/schisto/>
Plate: MS1-0095P-V361 row: 3 column: H.
Location/Qualifiers

FEATURES

```

1 494
source      /organism="Schistosoma mansoni"
            /mgi_type="mRNA"
            /db_xref="taxon:6183"
            /clone="MS1-0095P-V361-H03.B"
            /sex="mixed pool"
            /dev_stage="schistosomulum"
            /lab_host="in vitro culture"
            /clone_lib="MS1-0095"

```

NOTES

Query Match 11.2%; Score 46.2; DB 6; Length 494;
 Best Local Similarity 57.1%; Pred. No. 0.27;
 Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 48 TGAAGTACCGATGGCTGTGAACCTCGAAAGGTTCCGGAAACAAAAAGAGAGTGGTAGG 107
 |||||
 Db 233 TAAAGAATGTATCAATATAAACTTAACACCTCGAAATGAATATCAATCACAAAATGGTTGA 292
 |||||

Qy 108 TAATAATGTTTAATAATAGAAAAATAATAATAGTGGTGAAGAGGTTTGAAGTTGAGGA 167
 |||||
 Db 293 TGAATAATGTCACATAGGATGTTGATGATGATGGGACGATGATGACACGATGACGA 352
 |||||

Qy 168 AATTGAGGATAATGTAAGTATGATGACGA 194
 |||||
 Db 353 TGATGATGATAATGTGAATGATGAGA 379
 |||||

RESULT 12
 BM274859 365 bp mRNA linear EST 20-DEC-2001
 LOCUS PEST0aa75h03.y1 Plasmodium falciparum 3D7 gametocyte cDNA library
 DEFINITION Plasmodium falciparum 3D7 cDNA 5', mRNA sequence.

ACCESSION BM274859
 VERSION BM274859
 KEYWORDS EST.
 SOURCE Plasmodium falciparum 3D7
 ORGANISM Plasmodium falciparum 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 365)
 Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Pape, D.,
 Maria, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I.,
 Teagareishvili, R., Belaygorod, L., Franklin, C., Carr, L., Grow, A.,
 Maguire, L., Richey, J., Wadkins, J., Kennedy, S., Levinso, D.,
 Waterston, R., Wilson, R. and Sibley, D.

TITLE WashU Plasmodium EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: L. David Sibley
 WashU Plasmodium EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Library was constructed by R. Haywood. DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: L. David Sibley
 (sibley@borcim.wustl.edu), Washington University
 Seg primer: -40UP from Gibco
 High quality sequence stop: 345.

FEATURES
 source Location/Qualifiers
 1..365
 /organism="Plasmodium falciparum 3D7"
 /mol_type="mRNA"
 /db_xref="taxon:36329"
 /dev_stage="gametocyte (stage III-V)"
 /lab_hosts="DH10B (GeneHog, Invitrogen, Inc.)"
 /clone_lib="Plasmodium falciparum 3D7 gametocyte cDNA library"

/note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2: XhoI. The library was constructed by R. Haywood. cDNAs were synthesized from gametocyte poly(A)+ RNA by oligo d(T) priming, size-selected and directionally cloned into the EcoRI (5' end) to XhoI (3' end) sites of the Uni-ZAP XR lambda vector (Stratagene). The primary library was mass excised as phagemid using the ExAssist helper phage (Stratagene). Clones were mass excised using the ExAssist helper phage (Stratagene), the phagemids were precipitated with PEG 8000 and extracted with phenol/chloroform. Phagemid DNA was electroporated into DH10B cells. Clone Availability: David Sibley, Washington University."

ORIGIN

Query Match 11.1%; Score 45.6; DB 4; Length 365;
 Best Local Similarity 54.9%; Pred. No. 0.37;
 Matches 90; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 31 GTTGATGAGTTCGTAGATGAAGTACCGATGCTGTGAACTCGAAAGGTTCCGGAAACA 90
 |||||
 Db 7 GATGCGATATTATAGCTCAAGGTTCAAGTGGTCAACACACAAGAATATATAAAAGACGTA 66
 |||||

Qy 91 AAAAAAGAGAGCTGGTGTAGTAAATGTTTAATAAAGAAAAATAATAATAATAGTGGTAAAGAAG 150
 |||||
 Db 67 GAAATAATGTGAACACGAAATAATGAAAAACAATATTAAAAAAAAGAATATTGATAATAAT 126
 |||||

Qy 151 GGTTTGAAAGCTTGAGGAAATTTGAGGATAATCTTAAGTGTATGACGA 194
 |||||
 Db 127 GATAATAATGATAATAATGACGACGACGATGATGATGATGA 170
 |||||

RESULT 13
 BJ404348/c

LOCUS BJ404348 645 bp mRNA linear EST 10-MAR-2002
 DEFINITION BJ404348 Dictyostelium discoideum cDNA library, SF Dictyostelium discoideum cDNA clone dds28b07 3', mRNA sequence.

ACCESSION BJ404348
 VERSION BJ404348
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 REFERENCE 1 (bases 1 to 645)
 Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
 Full length cDNA of Dictyostelium discoideum at the slug stage
 Unpublished (2002)
 COMMENT Contact: Tadasu Shin-i
 Center for Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES
 source Location/Qualifiers

1..645
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="dds28b07"
 /sex="mat A"
 /dev_stage="slug stage"
 /clone_lib="Dictyostelium discoideum cDNA library, SF"

ORIGIN

Query Match 11.0%; Score 45.4; DB 4; Length 645;
 Best Local Similarity 59.8%; Pred. No. 0.45;
 Matches 76; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 68 AACTCGAAGGTTCCCGAAAAACAAAAGAGAGCTGGTAGCTTAATGTTTAATAAAGA 127
 |||||
 Db 584 AAATCGTATTGAATGCAACGAGAAAAATGAGCAAGTATCAAAATAATAATAATAATG 525
 |||||

Qy 128 AAATAATAATAAGTGGTGAAGAGGGTTTGAAGTTGAGGAAATGAGGATAATGTAAGTG 187
 |||||
 Db 524 ATATAATAATAATAGTAAATAATAATCAATAATGATAATGATAATGATATGATTG 465
 |||||

Qy 188 ATGACGA 194
 |||||
 Db 464 ATAATGA 458
 |||||

RESULT 14

CB552291 726 bp mRNA linear EST 01-JUN-2003
 LOCUS CB552291
 DEFINITION MMSP0051_D06 MMSP Macaca mulatta cDNA, mRNA sequence.
 ACCESSION CB552291

VERSION CB552291.1 GI:31301486
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 726)
AUTHORS Katze, M.G., Bumgarner, R., Korth, M., Feldman, R., Anjaji, M. and Holzman, T.
TITLE Expressed sequence tags from Rhesus macaque spleen
JOURNAL Unpublished (2002)
COMMENT Contact: Holzman T
Katze Lab
University of Washington
Box 358070, Seattle, WA 98195-8070, USA
Tel: 206 732 6156
Fax: 206 732 6055
Email: ted@locke.hs.washington.edu
Similar to GenBank entry HSU93049 U93049 Human SLP-76 associated protein mRNA, complete cds. 5/1997
Plate: MWSP0051 row: D column: 06.
FEATURES
source Location/Qualifiers
1..726
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/sex="male"
/cell_type="monocytes"
/dev_stage="adult"
/clone_lib="MWSP"
/note="Organ: spleen"
ORIGIN
Query Match 11.0%; Score 45.4; DB 6; Length 726;
Best Local Similarity 48.3%; Pred. No. 0.46;
Matches 127; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
Qy 10 ATTGAACCTCACTGAAAAAGTTGTTGATGAGTTCGTAGATGAAGTACCGATGGCTGTGAAA 69
Db 278 AATGAAGCAATCAAGATGGTGTCTATGCACTCTGATGGTCTGGAACCTAGATGAGAA 337
Qy 70 CTCGAAAGGTTCCGGAACAAAAAGAGAGTGGTAGGTAATAATGTTTAATAAAGAAA 129
Db 338 CAAGAGAGTGAAGGAGAAACATATGACACATAGAACATCCAAAGAAAGAGAGAGAAA 397
Qy 130 ATAAATATAGTGTGAAGAGGTTTGAAAGTTGAGGAAATGAGGATAATGTAAGTAT 189
Db 398 AGAGAAAAGGAGGAAAAGAGAGGTTAGAGCTGGAGAAAAAGGAAACAGAAAAGAGAA 457
Qy 190 GACGAGTCTATCGCTCATCGAGTACGTTTAAATCAATATCGCTTATACAACTCACTC 249
Db 458 AAGAAAGAACAGAAATAAAGAAATTTAAACTTAACAGGCCCTATTCAAGTCATCCAT 517
Qy 250 CGAGCCAAATTTGTTTACTTAAAT 272
Db 518 CTTGCAAAAGCTTGTGTGATG 540
RESULT 15
CV527869/c
LOCUS 729 bp mRNA linear EST 07-OCT-2004
DEFINITION CS_GIL_16E06.SP6 Blue crab gill, normalized Callinectes sapidus cDNA clone CS_GIL_16E06.5, similar to gb|AAH01555.1| PPIG protein - Homo sapiens. Score = 41.6 bits (96), Expect = 0.017, mRNA sequence.
ACCESSION CV527869
VERSION CV527869.1 GI:53912248
KEYWORDS EST.
SOURCE Callinectes sapidus (blue crab)
ORGANISM Callinectes sapidus
Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

Rubrachyura; Portunioidea; Portunidae; Callinectes.
1 (bases 1 to 729)
AUTHORS Shafer, T.H., Coblentz, F.E. and Towle, D.W.
TITLE Expressed sequence tags from normalized cDNA libraries prepared from gill and hypodermis tissues of the blue crab, Callinectes sapidus
JOURNAL Unpublished (2004)
COMMENT Contact: Thomas H. Shafer
Department of Biological Sciences
University of North Carolina Wilmington
601 S. College Rd, Wilmington, NC 28403, USA
Tel: 910-962-7275
Fax: 910-962-4066
Email: shafert@uncw.edu
Plate: 16 row: E column: 06
Seq primer: SP6
High quality sequence start: 3
High quality sequence stop: 515.
FEATURES
source Location/Qualifiers
1..729
/organism="Callinectes sapidus"
/mol_type="mRNA"
/db_xref="taxon:6763"
/clone="CS_GIL_16E06"
/tissue_type="Pooled anterior and posterior gills from crabs acclimated to salinities of 35 and 5 parts per thousand"
/dev_stage="Adult intermolt"
/clone_lib="Blue crab gill, normalized"
/note="Vector: PCMV Sport 6.1; Total RNA samples were prepared individually from each tissue, checked for quality, and then pooled for construction and normalization of a cDNA library by Invitrogen. Plasmids were isolated and inserts sequenced from their 5'-ends by the Blue Crab Molecular Genetics Laboratory at the University of North Carolina Wilmington. Traces were trimmed, compared (BLASTx) to NCBI non-redundant protein database as of 19 July 2004, and processed for submission to dbEST by trace2dbEST software (Parkinson, Anthony and Blaxter, unpublished software)."
ORIGIN
Query Match 11.0%; Score 45.2; DB 7; Length 729;
Best Local Similarity 58.0%; Pred. No. 0.52;
Matches 80; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 108 TAATAATGTTTAATAAAGAAAATAATAATAGTGGTAAGAGGTTTGAAAGTTGAGGA 167
Db 644 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATATGA 585
Qy 168 AATTGAGGATAATGTAAGTGAAGTATCGAGTCTATCGCTCATCGAGTACGTTTAAATCAAT 227
Db 584 TGATGATGATGATACATTGTTGATCTTAACAAATCAGGATACAAAAAGCTACATAAAT 525
Qy 228 ATGCCTTTATACAATCAAC 245
Db 524 ACAAATAATCCATCAAC 507
Search completed: September 26, 2005, 04:37:49
Job time : 3333.26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 01:18:09 ; Search time 160.434 Seconds
(without alignments)
4191.810 Million cell updates/sec

Title: US-09-551-494-12

Perfect score: 411

Sequence: 1 ccctcgcaattgaactcac.....tgtgcctagtacagatga 411

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCURS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336.2	81.8	7685	3	US-09-502-710-22
2	336.2	81.8	7685	3	US-09-502-710-25
3	336.2	81.8	7685	3	US-09-502-711-22
4	336.2	81.8	7685	3	US-09-502-711-25
5	336.2	81.8	7685	4	US-09-565-616A-1
6	336.2	81.8	7686	3	US-09-502-710-23
7	336.2	81.8	7686	3	US-09-502-710-26
8	336.2	81.8	7686	3	US-09-502-711-23
9	336.2	81.8	7686	3	US-09-502-711-26
10	336.2	81.8	7686	4	US-09-565-616A-2
11	336.2	81.8	7687	3	US-09-502-710-24
12	336.2	81.8	7687	3	US-09-502-711-24
13	336.2	81.8	7688	3	US-09-502-710-27
14	336.2	81.8	7688	3	US-09-502-711-27
15	36.4	23.5	7926	3	US-09-500-554-1
16	96.4	23.5	7926	3	US-09-726-648-1
17	96.4	23.5	7926	4	US-10-119-330-1
18	92.6	22.5	2621	2	US-08-553-619B-8
19	89.6	21.8	1860	2	US-08-553-619B-1
20	87.6	21.3	6395	2	US-08-687-559-2
21	87.6	21.3	6395	3	US-09-259-741-1
22	87.6	21.3	6395	3	US-09-037-751-1
23	87.6	21.3	6395	3	US-09-466-422-1
24	87.6	21.3	6395	4	US-09-401-415-2
25	87.6	21.3	6395	4	US-09-962-527-1
26	87.6	21.3	6425	3	US-09-259-741-3
27	87.6	21.3	6425	3	US-09-037-751-3

28	87.6	21.3	6425	3	US-09-466-422-3	Sequence 3, Appli
29	87.6	21.3	6425	4	US-09-962-527-3	Sequence 3, Appli
30	87.6	21.3	6439	3	US-09-259-741-2	Sequence 2, Appli
31	87.6	21.3	6439	3	US-09-037-751-2	Sequence 2, Appli
32	87.6	21.3	6439	3	US-09-466-422-2	Sequence 2, Appli
33	87.6	21.3	6439	4	US-09-962-527-2	Sequence 2, Appli
34	87.6	21.3	6446	3	US-09-259-741-5	Sequence 5, Appli
35	87.6	21.3	6446	3	US-09-037-751-5	Sequence 5, Appli
36	87.6	21.3	6446	3	US-09-466-422-5	Sequence 5, Appli
37	87.6	21.3	6446	4	US-09-962-527-5	Sequence 5, Appli
38	87.6	21.3	6475	3	US-09-259-741-4	Sequence 4, Appli
39	87.6	21.3	6475	3	US-09-037-751-4	Sequence 4, Appli
40	87.6	21.3	6475	3	US-09-466-422-4	Sequence 4, Appli
41	87.6	21.3	6475	4	US-09-962-527-4	Sequence 4, Appli
42	78.2	19.0	1561	4	US-09-419-788-24	Sequence 24, Appli
43	78.2	19.0	1582	4	US-09-419-788-25	Sequence 25, Appli
44	76.8	18.7	468	2	US-08-324-003A-26	Sequence 26, Appli
45	76.8	18.7	468	4	US-09-755-836-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1

US-09-502-710-22

; Sequence 22, Application US/09502710

; Patent No. 6300133

; GENERAL INFORMATION:

; APPLICANT: LINDBO, John A.

; APPLICANT: POGUE, Gregory P.

; APPLICANT: TURPEN, Thomas H.

; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM

; TITLE OF INVENTION: AN UNCAPPED SINGLE-COMPONENT RNA VIRUS

; FILE REFERENCE: 008010137US13

; CURRENT APPLICATION NUMBER: US/09/502,710

; CURRENT FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: 09/359,301

; PRIOR FILING DATE: 1999-07-21

; PRIOR APPLICATION NUMBER: 09/359,305

; PRIOR FILING DATE: 1999-07-21

; PRIOR APPLICATION NUMBER: 09/232,170

; PRIOR FILING DATE: 1999-01-15

; PRIOR APPLICATION NUMBER: 09/008,186

; PRIOR FILING DATE: 1998-01-16

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 22

; LENGTH: 7685

; TYPE: DNA

; ORGANISM: Tobacco mosaic virus

US-09-502-710-22

Query Match 81.8%; Score 336.2; DB 3; Length 7685;

Best Local Similarity 98.9%; Pred. No. 1.9e-87;

Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy	60	GGCTGTGAAACTCG-AAAGGTTCCGGAACACAAAAAGAGAGTGGTAAATAGTTA	118
Db	6681	GGCTGTGAAACTCGAAGGTTCCGGAACACAAAAAGAGAGTGGTAAATAGTTA	6740
Qy	119	ATAATAGAAAAATAATAATAGTGGTAAAGGGTTTGAAGTTGAGGAATTCAGGATA	178
Db	6741	ATAATAGAAAAATAATAATAGTGGTAAAGGGTTTGAAGTTGAGGAATTCAGGATA	6800
Qy	179	ATGTAAGTGTACGAGTCTATCGGTCATCGAGTACGTTTAACTTAATCAATATCCCTTATAC	238
Db	6801	ATGTAAGTGTACGAGTCTATCGGTCATCGAGTACGTTTAACTTAATCAATATCCCTTATAC	6860
Qy	239	AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTTCCGCTTATGCAGATCTCTGTGCAGCT	298
Db	6861	AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTTCCGCTTATGCAGATCTCTGTGCAGCT	6920
Qy	299	GATCAATCTGTGTACAAATGATTAAGTAACCAAGTTTCAACCAACCAAGTAGGACAAAC	358

```
Db 6921 GATCAATCTGTGTCAAAATGCTGGTAAACAGCTTTCAAAACGCAACAGCTAGGACAAC 6980
Qy 359 AGTCCAAACAGCAAAATTTGGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAAAATTTGGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 2
US-09-502-710-25
; Sequence 25, Application US/09502710
; Patent No. 6300134
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS
; FILE REFERENCE: 008010137US13
; CURRENT APPLICATION NUMBER: US/09/502,710
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-502-710-25

Query Match 81.8%; Score 336.2; DB 3; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAAACCAAAAGAGAGTGGTAGGTAATAGTTA 118
Db 6681 GCCTGTGAAACTCGAAAGGTTCCGGAAACCAAAAGAGAGTGGTAGGTAATAGTTA 6740
Qy 119 ATAATAGAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTCAGGATA 178
Db 6741 ATAATAGAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTCAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTCAAAATGATGAGTAAAGGGTTTGAAGTTGAGGAAATTCAGGATA 358
Db 6921 GATCAATCTGTGTCAAAATGATGAGTAAAGGGTTTGAAGTTGAGGAAATTCAGGATA 6980
Qy 359 AGTCCAAACAGCAAAATTTGGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAAAATTTGGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 3
US-09-502-711-22
; Sequence 22, Application US/09502711
; Patent No. 6300134
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
```

```
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-502-711-22

Query Match 81.8%; Score 336.2; DB 3; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 342; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAAACCAAAAGAGAGTGGTAGGTAATAGTTA 118
Db 6681 GCCTGTGAAACTCGAAAGGTTCCGGAAACCAAAAGAGAGTGGTAGGTAATAGTTA 6740
Qy 119 ATAATAGAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTCAGGATA 178
Db 6741 ATAATAGAAAATAAATAATAGTGGTAAGAGGGTTTGAAGTTGAGGAAATTCAGGATA 6800
Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTCAAAATGATGAGTAAAGGGTTTGAAGTTGAGGAAATTCAGGATA 358
Db 6921 GATCAATCTGTGTCAAAATGATGAGTAAAGGGTTTGAAGTTGAGGAAATTCAGGATA 6980
Qy 359 AGTCCAAACAGCAAAATTTGGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAAAATTTGGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 4
US-09-502-711-25
; Sequence 25, Application US/09502711
; Patent No. 6300134
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
```

```
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-502-711-25

Query Match      81.8%; Score 336.2; DB 3; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGCTAGGTAATAGTTA 118
Db 6681 GCCTGTGAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGCTAGGTAATAGTTA 6740

Qy 119 ATAATAAGAAATAAATAATAGTGTGAAGAGGTTTGAAGTTGAGGAAATGAGGATA 178
Db 6741 ATAATAAGAAATAAATAATAGTGTGAAGAGGTTTGAAGTTGAGGAAATGAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGAGATCCTGTGCAGCT 6920

Qy 299 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAAACGCAACAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAAACGCAACAGCTAGGACAAC 6980

Qy 359 AGTCAACAGCAATTTGGGATGCTGGAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCAACAGCAATTTGGGATGCTGGAACCTGTGCTAGTATGACAGTGA 7033

RESULT 5
US-09-565-616A-1
; Sequence 1, Application US/09565616A
; Patent No. 6656726
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P.
; APPLICANT: Pogue, Gregory P.
; APPLICANT: Lindbo, John A.
; TITLE OF INVENTION: VIRAL EXPRESSION VECTORS
; FILE REFERENCE: 008010165US01
; CURRENT APPLICATION NUMBER: US/09/565,616A
; CURRENT FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60,132,697
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-09-565-616A-1

Query Match      81.8%; Score 336.2; DB 4; Length 7685;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGCTAGGTAATAGTTA 118
Db 6681 GCCTGTGAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGCTAGGTAATAGTTA 6740

Qy 119 ATAATAAGAAATAAATAATAGTGTGAAGAGGTTTGAAGTTGAGGAAATGAGGATA 178
Db 6741 ATAATAAGAAATAAATAATAGTGTGAAGAGGTTTGAAGTTGAGGAAATGAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 7686
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-502-710-23

Query Match      81.8%; Score 336.2; DB 3; Length 7686;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGCTAGGTAATAGTTA 118
Db 6681 GCCTGTGAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGCTAGGTAATAGTTA 6740

Qy 119 ATAATAAGAAATAAATAATAGTGTGAAGAGGTTTGAAGTTGAGGAAATGAGGATA 178
Db 6741 ATAATAAGAAATAAATAATAGTGTGAAGAGGTTTGAAGTTGAGGAAATGAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATGAGATCCTGTGCAGCT 6920

Qy 299 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAAACGCAACAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGATAGTAAACAGTTTCAAACGCAACAGCTAGGACAAC 6980

Qy 359 AGTCAACAGCAATTTGGGATGCTGGAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCAACAGCAATTTGGGATGCTGGAACCTGTGCTAGTATGACAGTGA 7033

RESULT 6
US-09-502-710-23
; Sequence 23, Application US/09502710
; Patent No. 6300133
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; FILE REFERENCE: 008010137US13
; CURRENT APPLICATION NUMBER: US/09/502,710
; CURRENT FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 7686
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-502-710-23
```



```
US-09-502-711-26
Query Match      81.8%; Score 336.2; DB 3; Length 7686;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAACTCG-AAAGGTTCCGGAACCAAAAAGAGAGTGGTAGGTAATATGTTA 118
Db 6681 GCCTGTGAACTCGAAAGGTTCCGGAACCAAAAAGAGAGTGGTAGGTAATATGTTA 6740

Qy 119 ATAATAAGAAAATAAATAATAGTGTGTAAGAGGGTTTCGAAAGTTGAGGAAATGAGGATA 178
Db 6741 ATAATAAGAAAATAAATAATAGTGTGTAAGAGGGTTTCGAAAGTTGAGGAAATGAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGCGTCAATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGCGTCAATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 6860

Qy 239 AATCAACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCCTGTCAGCT 298
Db 6861 AATCAACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCCTGTCAGCT 6920

Qy 299 GATCAATCTGTGTAACAATGATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 358
Db 6921 GATCAATCTGTGTAACAATGATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 6980

Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGGAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGGAACCTGTGCTAGTATGACAGTGA 7033

RESULT 10
US-09-565-616A-2
; Sequence 2, Application US/09565616A
; Patent No. 6556726
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P.
; APPLICANT: Pogue, Gregory P.
; APPLICANT: Lindbo, John A.
; TITLE OF INVENTION: VIRAL EXPRESSION VECTORS
; FILE REFERENCE: 008010165US01
; CURRENT APPLICATION NUMBER: US/09/565,616A
; CURRENT FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60,132,697
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 7686
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-09-565-616A-2

Query Match      81.8%; Score 336.2; DB 4; Length 7686;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAACTCG-AAAGGTTCCGGAACCAAAAAGAGAGTGGTAGGTAATATGTTA 118
Db 6681 GCCTGTGAACTCGAAAGGTTCCGGAACCAAAAAGAGAGTGGTAGGTAATATGTTA 6740

Qy 119 ATAATAAGAAAATAAATAATAGTGTGTAAGAGGGTTTCGAAAGTTGAGGAAATGAGGATA 178
Db 6741 ATAATAAGAAAATAAATAATAGTGTGTAAGAGGGTTTCGAAAGTTGAGGAAATGAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGCGTCAATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGCGTCAATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 6860

Qy 239 AATCAACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCCTGTCAGCT 298
Db 6861 AATCAACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCCTGTCAGCT 6920

US-09-502-710-24
; Sequence 24, Application US/09502710
; Patent No. 6300133
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: AN UNCAPPED SINGLE-COMPONENT RNA VIRUS
; FILE REFERENCE: 008010137US13
; CURRENT APPLICATION NUMBER: US/09/502,710
; CURRENT FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 7687
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-502-710-24

Query Match      81.8%; Score 336.2; DB 3; Length 7687;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAACTCG-AAAGGTTCCGGAACCAAAAAGAGAGTGGTAGGTAATATGTTA 118
Db 6682 GCCTGTGAACTCGAAAGGTTCCGGAACCAAAAAGAGAGTGGTAGGTAATATGTTA 6741

Qy 119 ATAATAAGAAAATAAATAATAGTGTGTAAGAGGGTTTGAAGTTGAGGAAATGAGGATA 178
Db 6742 ATAATAAGAAAATAAATAATAGTGTGTAAGAGGGTTTGAAGTTGAGGAAATGAGGATA 6801

Qy 179 ATGTAAGTGATGACGAGTCTATCGCGTCAATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 238
Db 6802 ATGTAAGTGATGACGAGTCTATCGCGTCAATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 6861

Qy 239 AATCAACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCCTGTCAGCT 298
Db 6862 AATCAACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCCCTGTCAGCT 6921

Qy 299 GATCAATCTGTGTAACAATGATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 358
Db 6922 GATCAATCTGTGTAACAATGATAGGTAACAGTTTCAAAAGCAACAGCTAGGATA 6981

Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6982 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7034

RESULT 12
US-09-502-711-24
; Sequence 24, Application US/09502711
; Patent No. 6300134
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
```

```
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; TITLE OF INVENTION: SEQUENCE BETWEEN THE CAP AND 5' END
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/502,711
; CURRENT FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 7687
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
; US-09-502-711-24

Query Match      81.8%; Score 336.2; DB 3; Length 7687;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 118
Db 6682 GCCTGTGAAACTCGAAAGGTTCCGGAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 6741

Qy 119 ATAATAGAAATAAATAATAGTGGTAAGAGGGCTTTGAAAGTTGAGGAAATTCAGGATA 178
Db 6742 ATAATAGAAATAAATAATAGTGGTAAGAGGGCTTTGAAAGTTGAGGAAATTCAGGATA 6801

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCGTATAC 238
Db 6802 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCGTATAC 6861

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
Db 6862 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6921

Qy 299 GATCAATCTGTGTAACAATGATAGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 358
Db 6922 GATCAATCTGTGTAACAATGATAGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 6981

Qy 359 AGTCCAAACAGCAATTTCCGGATGCTGGAACCTGTGCTAGTATGACAGTGA 411
Db 6982 AGTCCAAACAGCAATTTCCGGATGCTGGAACCTGTGCTAGTATGACAGTGA 7034

RESULT 13
US-09-502-710-27
; Sequence 27, Application US/09502710
; Patent No. 6300133
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; TITLE OF INVENTION: SEQUENCE BETWEEN THE CAP AND 5' END
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/502,711
; CURRENT FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 7688
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
; US-09-502-711-27

Query Match      81.8%; Score 336.2; DB 3; Length 7688;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 118
Db 6684 GCCTGTGAAACTCGAAAGGTTCCGGAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 6743

Qy 119 ATAATAGAAATAAATAATAGTGGTAAGAGGGCTTTGAAAGTTGAGGAAATTCAGGATA 178
Db 6744 ATAATAGAAATAAATAATAGTGGTAAGAGGGCTTTGAAAGTTGAGGAAATTCAGGATA 6803

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCGTATAC 238
Db 6804 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCGTATAC 6863

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
Db 6864 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6923

Qy 299 GATCAATCTGTGTAACAATGATAGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 358
Db 6924 GATCAATCTGTGTAACAATGATAGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 6983

Qy 359 AGTCCAAACAGCAATTTCCGGATGCTGGAACCTGTGCTAGTATGACAGTGA 411
Db 6984 AGTCCAAACAGCAATTTCCGGATGCTGGAACCTGTGCTAGTATGACAGTGA 7036
```

```
RESULT 14
US-09-502-711-27
; Sequence 27, Application US/09502711
; Patent No. 6300134
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; TITLE OF INVENTION: SEQUENCE BETWEEN THE CAP AND 5' END
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/502,711
; CURRENT FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 7688
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
; US-09-502-711-27

Query Match      81.8%; Score 336.2; DB 3; Length 7688;
Best Local Similarity 98.9%; Pred. No. 1.9e-87;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 118
Db 6684 GCCTGTGAAACTCGAAAGGTTCCGGAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 6743

Qy 119 ATAATAGAAATAAATAATAGTGGTAAGAGGGCTTTGAAAGTTGAGGAAATTCAGGATA 178
Db 6744 ATAATAGAAATAAATAATAGTGGTAAGAGGGCTTTGAAAGTTGAGGAAATTCAGGATA 6803

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCGTATAC 238
Db 6804 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCGTATAC 6863

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
Db 6864 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6923

Qy 299 GATCAATCTGTGTAACAATGATAGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 358
Db 6924 GATCAATCTGTGTAACAATGATAGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 6983

Qy 359 AGTCCAAACAGCAATTTCCGGATGCTGGAACCTGTGCTAGTATGACAGTGA 411
Db 6984 AGTCCAAACAGCAATTTCCGGATGCTGGAACCTGTGCTAGTATGACAGTGA 7036
```

```

Db 6684 GCGTGTGAAACTCGAAAAGGTTCCGAAAAACAAAAAGAGAGTGGTAGGTAATAGTGTTA 6743
QY 119 ATAATAAGAAAATAAATAAGTAGTGAAGAGGTTTGAAAGTTGAGGAAATTCAGGATA 178
Db 6744 ATAATAAGAAAATAAATAAGTAGTGAAGAGGTTTGAAAGTTGAGGAAATTCAGGATA 6803
QY 179 ATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 238
Db 6804 ATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 6863
QY 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6864 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6923
QY 299 GATCAATCTGTGTACAAATGATAGGTAAACAGTTTCAAAACGCAACAAAGCTAGGACAAC 358
Db 6924 GATCAATCTGTGTACAAATGATAGGTAAACAGTTTCAAAACGCAACAAAGCTAGGACAAC 6983
QY 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6984 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7036

```

```

RESULT 15
US-09-500-554-1
; Sequence 1, Application US/09500554
; Patent No. 6284875
; GENERAL INFORMATION:
; APPLICANT: TURPEN, Thomas
; APPLICANT: GARGER, Stephen
; APPLICANT: MCCULLOCH, Michael
; APPLICANT: CAMERON, Terri
; APPLICANT: SAMNEK-POTTER, Michelle L.
; APPLICANT: HOLTZ, R. Barry
; TITLE OF INVENTION: METHOD FOR RECOVERING PROTEINS FROM THE
; TITLE OF INVENTION: INTERSTITIAL FLUID OF PLANT TISSUES
; FILE REFERENCE: 00801.0135.US00
; CURRENT APPLICATION NUMBER: US/09/500,554
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: 09/132,989
; PRIOR FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7926
; TYPE: DNA
; ORGANISM: VIRAL
US-09-500-554-1

```

```

Query Match 23.5%; Score 96.4; DB 3; Length 7926;
Best Local Similarity 63.2%; Pred. No. 8.4e-18;
Matches 148; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 178 AATGTAAGTGATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATA 237
Db 7194 ATTGAAGTAGAAGCGGAGACGCTCGGTCGGGATTCGATTCGTTAAATATATGCTTACT 7253
QY 238 CAATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGCAGATCCTGTGCAGC 297
Db 7254 CAATCAACTCTCCATCGCAATTTGTTTGTGTCATCTGTATGGGCTGACCCCTATAGAAT 7313
QY 298 TGATCAATCTGTGTACAAATGATAGGTAAACAGTTTCAAAACGCAACAAAGCTAGGACA 357
Db 7314 TGTAAAGTTTGTACAAATTTGTTAGGTAAACAGTTTCAAAACGCAACAAAGCAAGAACTA 7373
QY 358 CAGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 7374 CTGTTCAACAGCAGTTACGCGAGGTGTGGAACCTTCCCTCAGAGCACCGTCA 7427

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 01:30:34 ; Search time 724.658 Seconds
(without alignments)
3792.150 Million cell updates/sec

Title: US-09-551-494-12

Perfect score: 411

Sequence: 1 ccctgcgaattgaactcac.....tgctcctagtagacagtga 411

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	373.8	90.9	6355	15	US-10-321-434-7
2	336.2	81.8	7685	9	US-09-949-317-22
3	336.2	81.8	7685	9	US-09-949-317-25
4	336.2	81.8	7685	9	US-09-949-316-22
5	336.2	81.8	7685	9	US-09-949-316-25
6	336.2	81.8	7685	14	US-10-200-051-22
7	336.2	81.8	7685	14	US-10-200-051-25

8	336.2	81.8	7685	22	US-10-624-193-1	Sequence 1, Appl
9	336.2	81.8	7686	9	US-09-949-317-23	Sequence 23, Appl
10	336.2	81.8	7686	9	US-09-949-317-26	Sequence 26, Appl
11	336.2	81.8	7686	9	US-09-949-316-23	Sequence 23, Appl
12	336.2	81.8	7686	9	US-09-949-316-26	Sequence 26, Appl
13	336.2	81.8	7686	14	US-10-200-051-23	Sequence 23, Appl
14	336.2	81.8	7686	14	US-10-200-051-26	Sequence 26, Appl
15	336.2	81.8	7686	22	US-10-624-193-2	Sequence 2, Appl
16	336.2	81.8	7687	9	US-09-949-317-24	Sequence 24, Appl
17	336.2	81.8	7687	9	US-09-949-316-24	Sequence 24, Appl
18	336.2	81.8	7687	14	US-10-200-051-24	Sequence 24, Appl
19	336.2	81.8	7688	9	US-09-949-317-27	Sequence 27, Appl
20	336.2	81.8	7688	9	US-09-949-316-27	Sequence 27, Appl
21	336.2	81.8	7688	14	US-10-200-051-27	Sequence 27, Appl
22	336.2	81.8	8234	9	US-09-993-059-34	Sequence 34, Appl
23	336.2	81.8	8234	15	US-10-103-327-34	Sequence 34, Appl
24	336.2	81.8	8234	18	US-10-684-300-14	Sequence 14, Appl
25	336.2	81.8	8234	18	US-10-684-349-14	Sequence 14, Appl
26	336.2	81.8	8234	20	US-10-851-388-34	Sequence 34, Appl
27	336.2	81.8	8234	22	US-10-984-389-34	Sequence 34, Appl
28	336.2	81.8	10132	9	US-09-978-199-3	Sequence 3, Appl
29	336.2	81.8	10607	16	US-10-098-155-1	Sequence 1, Appl
30	336.2	81.8	10631	16	US-10-098-155-2	Sequence 2, Appl
31	328	79.8	2748	19	US-10-679-620-89	Sequence 89, Appl
32	328	79.8	2751	19	US-10-679-620-91	Sequence 91, Appl
33	328	79.8	10600	16	US-10-356-708-1	Sequence 1, Appl
34	328	79.8	10600	19	US-10-280-913A-1	Sequence 1, Appl
35	328	79.8	10600	19	US-10-684-134-1	Sequence 1, Appl
36	328	79.8	10600	19	US-10-637-758-1	Sequence 1, Appl
37	328	79.8	10624	16	US-10-356-708-2	Sequence 2, Appl
38	328	79.8	10624	19	US-10-280-913A-2	Sequence 2, Appl
39	328	79.8	10624	19	US-10-684-134-2	Sequence 2, Appl
40	328	79.8	10624	19	US-10-637-758-2	Sequence 2, Appl
41	328	79.8	11222	19	US-10-679-620-73	Sequence 73, Appl
42	210	51.1	769	15	US-10-211-079-21	Sequence 21, Appl
43	210	51.1	769	16	US-10-356-708-20	Sequence 20, Appl
44	210	51.1	769	19	US-10-280-913A-20	Sequence 20, Appl
45	210	51.1	769	19	US-10-684-134-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-10-321-434-7
; Sequence 7, Application US/10321434
; Publication No. US20030135882A1
; GENERAL INFORMATION:
; APPLICANT: Metzlaff, Michael
; APPLICANT: Meulewater, Frank
; APPLICANT: Gossel, Veronique
; APPLICANT: Fach, Ina
; TITLE OF INVENTION: Improved methods and means for delivering inhibitory RNA to plants
; FILE REFERENCE: PROMOD
; CURRENT APPLICATION NUMBER: US/10/321.434
; CURRENT FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 6355
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cdna sequence of the genome of TMV-U2
US-10-321-434-7

Query Match 90.9%; Score 373.8; DB 15; Length 6355;
Best Local Similarity 94.6%; Pred. No. 6.3e-84;
Matches 387; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 3 CTCGCCAATTGAACCTCACTGAAAAGTTGTGTGATTCGTAGATCGAAGTACCGATGCC 62
|||||

```
Db 5441 CTCGCCAATTGAACCTCACTGAAAAAGGTTGTTGAGGAGTTCGTGATGAAGTACCAATGGC 5500
Qy 63 TGTGAAACTCGAAAGGTTCCGGAACAAACAAAAAGAGAGTGTAGGTAAATATGTTAATAA 122
Db 5501 TGTGAACTCGAAAGGTTCCGGAACAAACAAAAAGAAATGGTAGTAAATATGTTAATAA 5560
Qy 123 TAAGAAAAATAAATAGTGTGAAGAAGGTTTGAAGTTGAGGAAATTTGAGGATAATGT 182
Db 5561 TAAGAAAAATAAATAACAGTGTGAAGAAGGTTTTAAAAATTTGAGGAAATTTGAGGATAATGT 5620
Qy 183 AAGTGATGACGAGTCTATCGCGTCAATCGAGTACGTTTTTAATCAATATGCCTTATACATC 242
Db 5621 AAGTGATGACGAGTCTATCGCGTCAATCGAGTACGTTTTTAATCAATATGCCTTATACATC 5680
Qy 243 AACTCTCGAGCCAAATTTGTTTACTTAAGTTCGCTTATGAGATCCTGTGCAGCTGATC 302
Db 5681 AACTCTCGAGCCAAATTTGTTTACTTAATCTTCCTGTTACGAGATCCTGTGCAGCTGATC 5740
Qy 303 AATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAAAAGCGCAACAAAGCTAGGACAAACAGTC 362
Db 5741 AATCTGTGTACAAATGCAATGGTAAACCAAGTTTCAAAAGCGCAACAAAGCTAGGACAAACAGTC 5800
Qy 363 CAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 5801 CAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 5849

RESULT 2
US-09-949-317-22
; Sequence 22, Application US/09949317
; Patent No. US20020164803A1
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; FILE REFERENCE: 008010137US13
; CURRENT APPLICATION NUMBER: US/09/949,317
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,710
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-949-317-22

Query Match 81.8%; Score 336.2; DB 9; Length 7685;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAACAAACAAAAAGAGAGTGTAGGTAATAGTTTA 118
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACAAACAAAAAGAGAGTGTAGGTAATAGTTTA 6740
Qy 119 ATAATAAGAAAAATAAATAAGTGTGAAGAAGGTTTGAAGTTTGAGGAAATTCAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAAGTGTGAAGAAGGTTTGAAGTTTGAGGAAATTCAGGATA 6800
Qy 179 ATGTAAGTATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAACACGCAACAAAGCTAGACAAAC 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAACACGCAACAAAGCTAGACAAAC 6980
Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 3
US-09-949-317-25
; Sequence 25, Application US/09949317
; Patent No. US20020164803A1
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; FILE REFERENCE: 008010137US13
; CURRENT APPLICATION NUMBER: US/09/949,317
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,710
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-949-317-25

Query Match 81.8%; Score 336.2; DB 9; Length 7685;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAACAAACAAAAAGAGAGTGTAGGTAATAGTTTA 118
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACAAACAAAAAGAGAGTGTAGGTAATAGTTTA 6740
Qy 119 ATAATAAGAAAAATAAATAAGTGTGAAGAAGGTTTGAAGTTTGAGGAAATTCAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAAGTGTGAAGAAGGTTTGAAGTTTGAGGAAATTCAGGATA 6800
Qy 179 ATGTAAGTATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAACACGCAACAAAGCTAGACAAAC 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAACACGCAACAAAGCTAGACAAAC 6980
Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033
```

```
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAACACGCAACAAAGCTAGACAAAC 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAACACGCAACAAAGCTAGACAAAC 6980
Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 3
US-09-949-317-25
; Sequence 25, Application US/09949317
; Patent No. US20020164803A1
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; FILE REFERENCE: 008010137US13
; CURRENT APPLICATION NUMBER: US/09/949,317
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,710
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-949-317-25

Query Match 81.8%; Score 336.2; DB 9; Length 7685;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAACAAACAAAAAGAGAGTGTAGGTAATAGTTTA 118
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACAAACAAAAAGAGAGTGTAGGTAATAGTTTA 6740
Qy 119 ATAATAAGAAAAATAAATAAGTGTGAAGAAGGTTTGAAGTTTGAGGAAATTCAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAAGTGTGAAGAAGGTTTGAAGTTTGAGGAAATTCAGGATA 6800
Qy 179 ATGTAAGTATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTATGACGAGTCTATCGCGTCATCGAGTACGTTTTTAATCAATATGCCTTATAC 6860
Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6920
Qy 299 GATCAATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAACACGCAACAAAGCTAGACAAAC 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGGTAACCAAGTTTCAACACGCAACAAAGCTAGACAAAC 6980
Qy 359 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033
```

```
RESULT 4
US-09-949-316-22
; Sequence 22, Application US/09949316
; Patent No. US20020168769A1
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: FOGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; TITLE OF INVENTION: SEQUENCE BETWEEN THE CAP AND 5' END
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/949,316
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-949-316-22

Query Match      81.8%; Score 336.2; DB 9; Length 7685;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAAAGAGAGTGGTAGGTAATAATGTTA 118
Db 6681 GCCTGTGAAACTCGGAAGGTTCCGGAACCAAAAAAGAGAGTGGTAGGTAATAATGTTA 6740

Qy 119 ATAATAAGAAAAATAAATAAGTGTGAAGAGGTTTGAAGTTGAGGAATTCAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAAGTGTGAAGAGGTTTGAAGTTGAGGAATTCAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920

Qy 299 GATCAATCTGTGTACAAATGCAATAGTAACTAGGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGTAACTAGGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 6980

Qy 359 AGTCCAAACAGCAATTTGGGATGCTCGAACCCTGCTAGTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGGGATGCTCGAACCCTGCTAGTAGTATGACAGTGA 7033

RESULT 5
US-09-949-316-25
; Sequence 25, Application US/09949316
; Patent No. US20020168769A1
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: FOGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; TITLE OF INVENTION: SEQUENCE BETWEEN THE CAP AND 5' END
```

```
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/949,316
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-09-949-316-25

Query Match      81.8%; Score 336.2; DB 9; Length 7685;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GCCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAAAGAGAGTGGTAGGTAATAATGTTA 118
Db 6681 GCCTGTGAAACTCGGAAGGTTCCGGAACCAAAAAAGAGAGTGGTAGGTAATAATGTTA 6740

Qy 119 ATAATAAGAAAAATAAATAAGTGTGAAGAGGTTTGAAGTTGAGGAATTCAGGATA 178
Db 6741 ATAATAAGAAAAATAAATAAGTGTGAAGAGGTTTGAAGTTGAGGAATTCAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920

Qy 299 GATCAATCTGTGTACAAATGCAATAGTAACTAGGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGCAATAGTAACTAGGTAACAGTTTCAAAACGCAACAGCTAGGACAAC 6980

Qy 359 AGTCCAAACAGCAATTTGGGATGCTCGAACCCTGCTAGTAGTATGACAGTGA 411
Db 6981 AGTCCAAACAGCAATTTGGGATGCTCGAACCCTGCTAGTAGTATGACAGTGA 7033

RESULT 6
US-10-200-051-22
; Sequence 22, Application US/10200051
; Publication No. US20030097683A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VECTORS DERIVED FROM A VIRUS AND CONTAINING
; TITLE OF INVENTION: INTERVENING SEQUENCE BETWEEN THE CAP AND THE 5' END AND ABLE TO I
; TITLE OF INVENTION: HOST PLANT CELL WITHIN A HOST PLANT
; FILE REFERENCE: 00801-0137-CN026
; CURRENT APPLICATION NUMBER: US/10/200,051
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 09/949,316
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
```

```
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-10-200-051-22

Query Match      81.8%; Score 336.2; DB 14; Length 7685;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      60  GGCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGTAGGTAATATGTTA 118
Db      6681 GGCTGTGAAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGTAGGTAATATGTTA 6740

Qy      119 ATAATAAGAAAAATAATAATAGTGTGAAGAGGTTTGAAGTTTGAGGAAATTGAGGATA 178
Db      6741 ATAATAAGAAAAATAATAATAGTGTGAAGAGGTTTGAAGTTTGAGGAAATTGAGGATA 6800

Qy      179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db      6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860

Qy      239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db      6861 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920

Qy      299 GATCAATCTGTGTACAAATGATTAAGTAAACCAAGTTTCAACGCAACAAAGCTAGGACAAC 358
Db      6921 GATCAATCTGTGTACAAATGATTAAGTAAACCAAGTTTCAACGCAACAAAGCTAGGACAAC 6980

Qy      359 AGTCCAAACAGCAATTTCCGGATGCTCGGAACCTGTGCCTAGTATGACAGTGA 411
Db      6981 AGTCCAAACAGCAATTTCCGGATGCTCGGAACCTGTGCCTAGTATGACAGTGA 7033

RESULT 7
US-10-200-051-25
; Sequence 25, Application US/10200051
; Publication No. US20030097683A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VECTORS DERIVED FROM A VIRUS AND CONTAINING
; TITLE OF INVENTION: INTERVENING SEQUENCE BETWEEN THE CAP AND THE 5' END AND ABLE TO
; TITLE OF INVENTION: HOST PLANT CELL WITHIN A HOST PLANT
; FILE REFERENCE: 00801-0137-CN026
; CURRENT APPLICATION NUMBER: US/10/200,051
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 09/949,316
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-10-200-051-25

Query Match      81.8%; Score 336.2; DB 14; Length 7685;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```

```
Qy      60  GGCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGTAGGTAATATGTTA 118
Db      6681 GGCTGTGAAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGTAGGTAATATGTTA 6740

Qy      119 ATAATAAGAAAAATAATAATAGTGTGAAGAGGTTTGAAGTTTGAGGAAATTGAGGATA 178
Db      6741 ATAATAAGAAAAATAATAATAGTGTGAAGAGGTTTGAAGTTTGAGGAAATTGAGGATA 6800

Qy      179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db      6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860

Qy      239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db      6861 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920

Qy      299 GATCAATCTGTGTACAAATGATTAAGTAAACCAAGTTTCAACGCAACAAAGCTAGGACAAC 358
Db      6921 GATCAATCTGTGTACAAATGATTAAGTAAACCAAGTTTCAACGCAACAAAGCTAGGACAAC 6980

Qy      359 AGTCCAAACAGCAATTTCCGGATGCTCGGAACCTGTGCCTAGTATGACAGTGA 411
Db      6981 AGTCCAAACAGCAATTTCCGGATGCTCGGAACCTGTGCCTAGTATGACAGTGA 7033

RESULT 8
US-10-624-193-1
; Sequence 1, Application US/10624193
; Publication No. US20050175590A1
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P
; APPLICANT: Pogue, Gregory P
; APPLICANT: Lindbo, John A
; TITLE OF INVENTION: VIRAL EXPRESSION VECTORS
; FILE REFERENCE: 60-016511US
; CURRENT APPLICATION NUMBER: US/10/624,193
; CURRENT FILING DATE: 2003-07-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7685
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-624-193-1

Query Match      81.8%; Score 336.2; DB 22; Length 7685;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      60  GGCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGTAGGTAATATGTTA 118
Db      6681 GGCTGTGAAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGTAGGTAATATGTTA 6740

Qy      119 ATAATAAGAAAAATAATAATAGTGTGAAGAGGTTTGAAGTTTGAGGAAATTGAGGATA 178
Db      6741 ATAATAAGAAAAATAATAATAGTGTGAAGAGGTTTGAAGTTTGAGGAAATTGAGGATA 6800

Qy      179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 238
Db      6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCCTTATAC 6860

Qy      239 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 298
Db      6861 AATCAACTCTCCGAGCCAAATTTGTTTAACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCT 6920

Qy      299 GATCAATCTGTGTACAAATGATTAAGTAAACCAAGTTTCAACGCAACAAAGCTAGGACAAC 358
Db      6921 GATCAATCTGTGTACAAATGATTAAGTAAACCAAGTTTCAACGCAACAAAGCTAGGACAAC 6980

Qy      359 AGTCCAAACAGCAATTTCCGGATGCTCGGAACCTGTGCCTAGTATGACAGTGA 411
Db      6981 AGTCCAAACAGCAATTTCCGGATGCTCGGAACCTGTGCCTAGTATGACAGTGA 7033
```

```

;
; CURRENT APPLICATION NUMBER: US/09/949,317
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,710
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 7686
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
;
US-09-949-317-26

Query Match      81.8%; Score 336.2; DB 9; Length 7686;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      60  GGCTGTGAAACTCG-AAAGGTTCCGGAAAAACAAAAAGAGAGTGTAGGTAATAGTTA 118
      |||
Db      6681  GGCTGTGAAACTCGAAAAAGTTCCGGAAAAACAAAAAGAGAGTGTAGGTAATAGTTA 6740

Qy      119  ATATAAGAAAAATAAATAAGTGTAAAGGGTTTGAAGTTGAGGAAATTCAGGATA 178
      |||
Db      6741  ATATAAGAAAAATAAATAAGTGTAAAGGGTTTGAAGTTGAGGAAATTCAGGATA 6800

Qy      179  ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
      |||
Db      6801  ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6920

Qy      239  AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
      |||
Db      6861  AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6920

Qy      299  GATCAATCTGTGTACAAATGATTAAGTAAACAGTTTCAAACGCAACAAAGCTAGGACAAC 358
      |||
Db      6921  GATCAATCTGTGTACAAATGATTAAGTAAACAGTTTCAAACGCAACAAAGCTAGGACAAC 6980

Qy      359  AGTCCAAACAGCAATTTCCGATGCTCGAAACCTGCTGCTAGTATGACAGTGA 411
      |||
Db      6981  AGTCCAAACAGCAATTTCCGATGCTCGAAACCTGCTGCTAGTATGACAGTGA 7033

RESULT 11
US-09-949-316-23
; Sequence 23, Application US/09949316
; Patent No. US20020168769A1
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VIRUS AND CONTAIN AN INTERVENING
; TITLE OF INVENTION: SEQUENCE BETWEEN THE CAP AND 5' END
; FILE REFERENCE: 008010137US15
; CURRENT APPLICATION NUMBER: US/09/949,316
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; FILE REFERENCE: 008010137US13
```

```

;
; CURRENT APPLICATION NUMBER: US/09/949,317
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,710
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 7686
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
;
US-09-949-317-23

Query Match      81.8%; Score 336.2; DB 9; Length 7686;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      60  GGCTGTGAAACTCG-AAAGGTTCCGGAAAAACAAAAAGAGAGTGTAGGTAATAGTTA 118
      |||
Db      6681  GGCTGTGAAACTCGAAAAAGTTCCGGAAAAACAAAAAGAGAGTGTAGGTAATAGTTA 6740

Qy      119  ATATAAGAAAAATAAATAAGTGTAAAGGGTTTGAAGTTGAGGAAATTCAGGATA 178
      |||
Db      6741  ATATAAGAAAAATAAATAAGTGTAAAGGGTTTGAAGTTGAGGAAATTCAGGATA 6800

Qy      179  ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
      |||
Db      6801  ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6920

Qy      239  AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATCGAGATCCTGTGCAGCT 298
      |||
Db      6861  AATCAACTCTCCGAGCCAAATTTGTTTAAAGTTCCGCTTATCGAGATCCTGTGCAGCT 6920

Qy      299  GATCAATCTGTGTACAAATGATTAAGTAAACAGTTTCAAACGCAACAAAGCTAGGACAAC 358
      |||
Db      6921  GATCAATCTGTGTACAAATGATTAAGTAAACAGTTTCAAACGCAACAAAGCTAGGACAAC 6980

Qy      359  AGTCCAAACAGCAATTTCCGATGCTCGAAACCTGCTGCTAGTATGACAGTGA 411
      |||
Db      6981  AGTCCAAACAGCAATTTCCGATGCTCGAAACCTGCTGCTAGTATGACAGTGA 7033

RESULT 10
US-09-949-317-26
; Sequence 26, Application US/09949317
; Patent No. US20020164803A1
; GENERAL INFORMATION:
; APPLICANT: LINDBO, John A.
; APPLICANT: POGUE, Gregory P.
; APPLICANT: TURPEN, Thomas H.
; TITLE OF INVENTION: RNA TRANSFORMATION VECTORS DERIVED FROM
; TITLE OF INVENTION: AN UNCAPPED SINGLE-COMPONENT RNA VIRUS
; FILE REFERENCE: 008010137US13
```



```
Qy 239 AATCAACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCTCTGTGCAGCT 298
Db 6861 AATCAACTCTCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCTCTGTGCAGCT 6920

Qy 299 GATCAATCTGTGTACAAATGATTAGTAAACAGTTTCAAAGCGCAACAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGATTAGTAAACAGTTTCAAAGCGCAACAGCTAGGACAAC 6980

Qy 359 AGTCCAACAGCAAAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAACAGCAAAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033

RESULT 14
US-10-200-051-26
; Sequence 26, Application US/10200051
; Publication No. US20030097683A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: A SINGLE-COMPONENT RNA VECTORS DERIVED FROM A VIRUS AND CONTAININ
; TITLE OF INVENTION: INTERVENING SEQUENCE BETWEEN THE CAP AND THE 5' END AND ABLE TO
; TITLE OF INVENTION: HOST PLANT CELL WITHIN A HOST PLANT
; FILE REFERENCE: 00801-0137-CN026
; CURRENT APPLICATION NUMBER: US/10/200,051
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 09/949,316
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/502,711
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/359,301
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/359,305
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/232,170
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09/008,186
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 7686
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus
US-10-200-051-26

Query Match 81.8%; Score 336.2; DB 14; Length 7686;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGGTAGGTAATAATGTTA 118
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGGTAGGTAATAATGTTA 6740

Qy 119 ATAATAAGAAAATAAATAATAGTGTGAAGAGGTTTGAAGGAAATTCAGGATA 178
Db 6741 ATAATAAGAAAATAAATAATAGTGTGAAGAGGTTTGAAGGAAATTCAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGACAGTGA 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGACAGTGA 6860

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCTCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCTCTGTGCAGCT 6920

Qy 299 GATCAATCTGTGTACAAATGATTAGTAAACAGTTTCAAAGCGCAACAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGATTAGTAAACAGTTTCAAAGCGCAACAGCTAGGACAAC 6980

Qy 359 AGTCCAACAGCAAAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAACAGCAAAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033
```

```
RESULT 15
US-10-624-193-2
; Sequence 2, Application US/10624193
; Publication No. US20050175590A1
; GENERAL INFORMATION:
; APPLICANT: Fitzmaurice, Wayne P
; APPLICANT: Pogue, Gregory P
; APPLICANT: Lindbo, John A
; TITLE OF INVENTION: VIRAL EXPRESSION VECTORS
; FILE REFERENCE: 60-016511US
; CURRENT APPLICATION NUMBER: US/10/624,193
; CURRENT FILING DATE: 2003-07-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 7686
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-624-193-2

Query Match 81.8%; Score 336.2; DB 22; Length 7686;
Best Local Similarity 98.9%; Pred. No. 2.2e-74;
Matches 349; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 60 GGCTGTGAAACTCG-AAAGGTTCCGGAACCAAAAGAGAGTGGTAGGTAATAATGTTA 118
Db 6681 GGCTGTGAAACTCGAAAGGTTCCGGAACCAAAAGAGAGTGGTAGGTAATAATGTTA 6740

Qy 119 ATAATAAGAAAATAAATAATAGTGTGAAGAGGTTTGAAGGAAATTCAGGATA 178
Db 6741 ATAATAAGAAAATAAATAATAGTGTGAAGAGGTTTGAAGGAAATTCAGGATA 6800

Qy 179 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGACAGTGA 238
Db 6801 ATGTAAGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGACAGTGA 6860

Qy 239 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCTCTGTGCAGCT 298
Db 6861 AATCAACTCTCCGAGCCAAATTTGTTTACTTAAGTTCCGCTTATGAGATCTCTGTGCAGCT 6920

Qy 299 GATCAATCTGTGTACAAATGATTAGTAAACAGTTTCAAAGCGCAACAGCTAGGACAAC 358
Db 6921 GATCAATCTGTGTACAAATGATTAGTAAACAGTTTCAAAGCGCAACAGCTAGGACAAC 6980

Qy 359 AGTCCAACAGCAAAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 411
Db 6981 AGTCCAACAGCAAAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGA 7033
```

Search completed: September 26, 2005, 05:02:40
Job time : 727.658 secs

This Page Blank (usp10)